



RN  
RN  
RP  
RC  
RX  
RA  
RA  
RT  
RT  
RL  
DR  
EMBL:  
DR  
KW  
SEQUENCE 1455 AA; 162914 MW; 7D729DAO CRC32;

[1]  
SEQUENCE FROM N.A.  
STRAIN-MT4239;  
MEDLINE; 98231643.  
FUJIWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S.,  
KIMURA S.; HAMADA S.;  
"Molecular analyses of glucosyltransferase genes among strains of  
Streptococcus mutans";  
FEMS MICROBIOL. LETT. 161:331-336(1998).  
EMBL; D88658; D1027050; -.

Db 578 VPSYSFIR-AHSEVQDLI 595  
SQ 1 VPSYSFIRAHSEVQDLI 19

Query Match 82.2%; Score 111; DB 2; Length 1455;  
Best Local Similarity 94.7%; Pred. No. 2.83e-12; 0; Mismatches 0; Indels 1; Gaps 1;  
Matches 18; Conservative

Query Match 82.2%; Score 111; DB 2; Length 1455;  
Best Local Similarity 94.7%; Pred. No. 2.83e-12; 0; Mismatches 0; Indels 1; Gaps 1;  
Matches 18; Conservative

RESULT 3 PRELIMINARY; PRT; 1455 AA.  
ID 059382  
AC 059382;  
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE GLUCOSYLTRANSFERASE-SI.  
GN GTFC.  
OS STREPTOCOCCUS MUTANS.  
OC BACTERIA; FIRMITICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
OC STREPTOCOCCUS.  
RN [1]  
SEQUENCE FROM N.A.

RA FUJIWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S.,  
RA KIMURA S.; HAMADA S.;  
RT "Molecular analyses of glucosyltransferase genes among strains of  
Streptococcus mutans";  
FEMS MICROBIOL. LETT. 161:331-336(1998).  
RL FEMS MICROBIOL. LETT. 161:331-336(1998).  
DR EMBL; D88652; D1027042; -.  
KW TRANSFERASE.  
SQ SEQUENCE 1455 AA; 162970 MW; 40B022BD CRC32;

Query Match 82.2%; Score 111; DB 2; Length 1455;  
Best Local Similarity 94.7%; Pred. No. 2.83e-12; 0; Mismatches 0; Indels 1; Gaps 1;  
Matches 18; Conservative

Db 578 VPSYSFIR-AHSEVQDLI 595  
SQ 1 VPSYSFIRAHSEVQDLI 19

Query Match 82.2%; Score 111; DB 2; Length 1455;  
Best Local Similarity 94.7%; Pred. No. 2.83e-12; 0; Mismatches 0; Indels 1; Gaps 1;  
Matches 18; Conservative

RESULT 4 PRELIMINARY; PRT; 1455 AA.  
ID 059397  
AC 059397;  
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE GLUCOSYLTRANSFERASE-SI.  
GN GTFC.  
OS STREPTOCOCCUS MUTANS.  
OC BACTERIA; FIRMITICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
OC STREPTOCOCCUS.  
RN [1]  
SEQUENCE FROM N.A.

RA FUJIWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S.,  
RA KIMURA S.; HAMADA S.;  
RT "Molecular analyses of glucosyltransferase genes among strains of  
Streptococcus mutans";  
FEMS MICROBIOL. LETT. 161:331-336(1998).  
RL FEMS MICROBIOL. LETT. 161:331-336(1998).  
DR EMBL; D88660; D1027053; -.  
KW TRANSFERASE.

Db 578 VPSYSFIR-AHSEVQDLI 595  
SQ 1 VPSYSFIRAHSEVQDLI 19

Query Match 82.2%; Score 111; DB 2; Length 1455;  
Best Local Similarity 94.7%; Pred. No. 2.83e-12; 0; Mismatches 0; Indels 1; Gaps 1;  
Matches 18; Conservative

RESULT 5 PRELIMINARY; PRT; 1455 AA.  
ID 069391  
AC 069391;  
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE GLUCOSYLTRANSFERASE-SI.  
GN GTFC.  
OS STREPTOCOCCUS MUTANS.  
OC BACTERIA; FIRMITICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
OC STREPTOCOCCUS.  
RN [1]  
SEQUENCE FROM N.A.

RA FUJIWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S.,  
RA KIMURA S.; HAMADA S.;  
RT "Molecular analyses of glucosyltransferase genes among strains of  
Streptococcus mutans";  
FEMS MICROBIOL. LETT. 161:331-336(1998).  
RL FEMS MICROBIOL. LETT. 161:331-336(1998).  
DR EMBL; D88660; D1027053; -.  
KW TRANSFERASE.

Db 578 VPSYSFIR-AHSEVQDLI 595  
SQ 1 VPSYSFIRAHSEVQDLI 19

Query Match 82.2%; Score 111; DB 2; Length 1455;  
Best Local Similarity 94.7%; Pred. No. 2.83e-12; 0; Mismatches 0; Indels 1; Gaps 1;  
Matches 18; Conservative

RESULT 6 PRELIMINARY; PRT; 1476 AA.  
ID 069390  
AC 069390;  
DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE GLUCOSYLTRANSFERASE-SI.  
GN GTFB.  
OS STREPTOCOCCUS MUTANS.  
OC BACTERIA; FIRMITICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
OC STREPTOCOCCUS.  
RN [1]  
SEQUENCE FROM N.A.

RA FUJIWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S.,  
RA KIMURA S.; HAMADA S.;  
RT "Molecular analyses of glucosyltransferase genes among strains of  
Streptococcus mutans";  
FEMS MICROBIOL. LETT. 161:331-336(1998).  
RL FEMS MICROBIOL. LETT. 161:331-336(1998).  
DR EMBL; D88660; D1027053; -.  
KW TRANSFERASE.

Db 578 VPSYSFIR-AHSEVQDLI 595  
SQ 1 VPSYSFIRAHSEVQDLI 19

Query Match 82.2%; Score 111; DB 2; Length 1476;

|                       |   |   |               |                       |   |   |      |                       |   |   |   |          |                       |   |   |      |          |
|-----------------------|---|---|---------------|-----------------------|---|---|------|-----------------------|---|---|---|----------|-----------------------|---|---|------|----------|
| Best Local Similarity | 94.7%   | Pred.                                   | No. 2.83e-12; | Matches               | 18;   | Conservative                            | 0;   | Mismatches            | 0;  | Indels                                  | 1;                                      | Gaps     | 1;                    |   |   |      |          |
| Db                    | 552   | VPSYSFIR-AHDSEVQDLI                     | 569           | ID                    | 069381;   | PRELIMINARY;                            | PRT; | 1476 AA.              | RESULT  | 9                                       | PRELIMINARY;                            | PRT;     | 1476 AA.              |   |   |      |          |
| Qy                    | 1   | VPSYSFIRAHDSEVQDLI                      | 19            | AC                    | 069381;   |   |      |                       | ID  | 069381;                                 | PRELIMINARY;                            | PRT;     | 1476 AA.              |   |   |      |          |
| DT                    | 01-AUG-1998   | (TREMBLREL. 07, CREATED)                |               | DT                    | 01-AUG-1998   | (TREMBLREL. 07, CREATED)                |      |                       | AC  | 069381;                                 |   |          |                       |   |   |      |          |
| DT                    | 01-AUG-1998   | (TREMBLREL. 07, LAST SEQUENCE UPDATE)   |               | DT                    | 01-AUG-1998   | (TREMBLREL. 07, LAST SEQUENCE UPDATE)   |      |                       | DT  | 01-AUG-1998                             | (TREMBLREL. 07, LAST SEQUENCE UPDATE)   |          |                       |   |   |      |          |
| DE                    | GLUCOSYLTRANSFERASE-I.  |   |               | DT                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DE  | GLUCOSYLTRANSFERASE-I.                  |   |          |                       |   |   |      |          |
| GN                    | GTFB.   |   |               | DR                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DR  | 01-NOV-1998                             | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |          |                       |   |   |      |          |
| OS                    | STREPTOCOCCUS MUTANS  |   |               | DR                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DR  | 01-NOV-1998                             | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |          |                       |   |   |      |          |
| OC                    | BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;   |   |               | DR                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DR  | 01-NOV-1998                             | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |          |                       |   |   |      |          |
| OC                    | STREPTOCOCCUS.  |   |               | DR                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DR  | 01-NOV-1998                             | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |          |                       |   |   |      |          |
| RN                    | [1]   |   |               | DR                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DR  | 01-NOV-1998                             | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |          |                       |   |   |      |          |
| RP                    | SEQUENCE FROM N.A.  |   |               | DR                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DR  | 01-NOV-1998                             | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |          |                       |   |   |      |          |
| RA                    | FUJIWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S., |   |               | DR                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DR  | 01-NOV-1998                             | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |          |                       |   |   |      |          |
| RA                    | KIMURA S., HAMADA S.;   |   |               | DR                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DR  | 01-NOV-1998                             | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |          |                       |   |   |      |          |
| RT                    | "Molecular analyses of glucosyltransferase genes among strains of     |   |               | DR                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DR  | 01-NOV-1998                             | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |          |                       |   |   |      |          |
| RT                    | Streptococcus mutans."  |   |               | DR                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DR  | 01-NOV-1998                             | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |          |                       |   |   |      |          |
| RL                    | FEMS MICROBIOL. LETT. 161:331-336(1998).                              |   |               | DR                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DR  | 01-NOV-1998                             | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |          |                       |   |   |      |          |
| DR                    | EMBL: D88557; D1027049; -.  |   |               | DR                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DR  | 01-NOV-1998                             | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |          |                       |   |   |      |          |
| KW                    | TRANSFERASE.  |   |               | DR                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DR  | 01-NOV-1998                             | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |          |                       |   |   |      |          |
| SQ                    | SEQUENCE 1476 AA; 165819 MW: AFGDBA6F CRC32;                          |   |               | DR                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DR  | 01-NOV-1998                             | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |          |                       |   |   |      |          |
| Query Match           | 82.2%; Score 111; DB 2; Length 1476;                                  |   |               | DR                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DR  | 01-NOV-1998                             | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |          |                       |   |   |      |          |
| Best Local Similarity | 94.7%; Pred. No. 2.83e-12;  |   |               | DR                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DR  | 01-NOV-1998                             | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |          |                       |   |   |      |          |
| Matches               | 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;                   |   |               | DR                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |                       | DR  | 01-NOV-1998                             | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |          |                       |   |   |      |          |
| Db                    | 552   | VPSYSFIR-AHDSEVQDLI                     | 569           | Qy                    | 1   | VPSYSFIRAHDSEVQDLI                      | 19   | RESULT                | 8   | PRELIMINARY;                            | PRT;                                    | 1476 AA. | RESULT                | 9   | PRELIMINARY;                            | PRT; | 1476 AA. |
| Qy                    |   |   |               | AC                    | 069384;   |   |      | AC                    | 069395;   |   |   |          | AC                    | 069395;   |   |      |          |
| DT                    | 01-AUG-1998   | (TREMBLREL. 07, CREATED)                |               | DT                    | 01-AUG-1998   | (TREMBLREL. 07, CREATED)                |      | DT                    | 01-AUG-1998   | (TREMBLREL. 07, CREATED)                |   |          | DT                    | 01-AUG-1998   | (TREMBLREL. 07, CREATED)                |      |          |
| DT                    | 01-AUG-1998   | (TREMBLREL. 07, LAST SEQUENCE UPDATE)   |               | DT                    | 01-AUG-1998   | (TREMBLREL. 07, LAST SEQUENCE UPDATE)   |      | DT                    | 01-AUG-1998   | (TREMBLREL. 07, LAST SEQUENCE UPDATE)   |   |          | DT                    | 01-AUG-1998   | (TREMBLREL. 07, LAST SEQUENCE UPDATE)   |      |          |
| DT                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |               | DT                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      | DT                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |   |          | DT                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |          |
| DE                    | GLUCOSYLTRANSFERASE-I.  |   |               | DE                    | GLUCOSYLTRANSFERASE-I.  |   |      | DE                    | GLUCOSYLTRANSFERASE-I.  |   |   |          | DE                    | GLUCOSYLTRANSFERASE-I.  |   |      |          |
| GN                    | GTFB.   |   |               | GN                    | GTFB.   |   |      | GN                    | GTFB.   |   |   |          | GN                    | GTFB.   |   |      |          |
| OS                    | STREPTOCOCCUS MUTANS  |   |               | OS                    | STREPTOCOCCUS MUTANS  |   |      | OS                    | STREPTOCOCCUS MUTANS  |   |   |          | OS                    | STREPTOCOCCUS MUTANS  |   |      |          |
| OC                    | BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;   |   |               | OC                    | BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;   |   |      | OC                    | BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;   |   |   |          | OC                    | BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;   |   |      |          |
| OC                    | STREPTOCOCCUS.  |   |               | OC                    | STREPTOCOCCUS.  |   |      | OC                    | STREPTOCOCCUS.  |   |   |          | OC                    | STREPTOCOCCUS.  |   |      |          |
| RN                    | [1]   |   |               | RN                    | [1]   |   |      | RN                    | [1]   |   |   |          | RN                    | [1]   |   |      |          |
| RP                    | SEQUENCE FROM N.A.  |   |               | RP                    | SEQUENCE FROM N.A.  |   |      | RP                    | SEQUENCE FROM N.A.  |   |   |          | RP                    | SEQUENCE FROM N.A.  |   |      |          |
| RC                    | STRAIN=MT4245;  |   |               | RC                    | STRAIN=MT4245;  |   |      | RC                    | STRAIN=MT4245;  |   |   |          | RC                    | STRAIN=MT4245;  |   |      |          |
| RX                    | STREPTOCOCCUS MUTANS  |   |               | RX                    | STREPTOCOCCUS MUTANS  |   |      | RX                    | STREPTOCOCCUS MUTANS  |   |   |          | RX                    | STREPTOCOCCUS MUTANS  |   |      |          |
| RA                    | FUJIWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S., |   |               | RA                    | FUJIWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S., |   |      | RA                    | FUJIWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S., |   |   |          | RA                    | FUJIWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S., |   |      |          |
| RA                    | KIMURA S., HAMADA S.;   |   |               | RA                    | KIMURA S., HAMADA S.;   |   |      | RA                    | KIMURA S., HAMADA S.;   |   |   |          | RA                    | KIMURA S., HAMADA S.;   |   |      |          |
| RT                    | "Molecular analyses of glucosyltransferase genes among strains of     |   |               | RT                    | "Molecular analyses of glucosyltransferase genes among strains of     |   |      | RT                    | "Molecular analyses of glucosyltransferase genes among strains of     |   |   |          | RT                    | "Molecular analyses of glucosyltransferase genes among strains of     |   |      |          |
| RT                    | Streptococcus mutans."  |   |               | RT                    | Streptococcus mutans."  |   |      | RT                    | Streptococcus mutans."  |   |   |          | RT                    | Streptococcus mutans."  |   |      |          |
| RL                    | FEMS MICROBIOL. LETT. 161:331-336(1998).                              |   |               | RL                    | FEMS MICROBIOL. LETT. 161:331-336(1998).                              |   |      | RL                    | FEMS MICROBIOL. LETT. 161:331-336(1998).                              |   |   |          | RL                    | FEMS MICROBIOL. LETT. 161:331-336(1998).                              |   |      |          |
| DR                    | EMBL: D88554; D1027045; -.  |   |               | DR                    | EMBL: D88554; D1027045; -.  |   |      | DR                    | EMBL: D88554; D1027045; -.  |   |   |          | DR                    | EMBL: D88554; D1027045; -.  |   |      |          |
| KW                    | TRANSFERASE.  |   |               | KW                    | TRANSFERASE.  |   |      | KW                    | TRANSFERASE.  |   |   |          | KW                    | TRANSFERASE.  |   |      |          |
| SQ                    | SEQUENCE 1476 AA; 165792 MW: C76D8483 CRC32;                          |   |               | SQ                    | SEQUENCE 1476 AA; 165647 MW: 258D4ADD CRC32;                          |   |      | SQ                    | SEQUENCE 1476 AA; 165647 MW: 258D4ADD CRC32;                          |   |   |          | SQ                    | SEQUENCE 1476 AA; 165647 MW: 258D4ADD CRC32;                          |   |      |          |
| Query Match           | 82.2%; Score 111; DB 2; Length 1476;                                  |   |               | Query Match           | 82.2%; Score 111; DB 2; Length 1476;                                  |   |      | Query Match           | 82.2%; Score 111; DB 2; Length 1476;                                  |   |   |          | Query Match           | 82.2%; Score 111; DB 2; Length 1476;                                  |   |      |          |
| Best Local Similarity | 94.7%; Pred. No. 2.83e-12;  |   |               | Best Local Similarity | 94.7%; Pred. No. 2.83e-12;  |   |      | Best Local Similarity | 94.7%; Pred. No. 2.83e-12;  |   |   |          | Best Local Similarity | 94.7%; Pred. No. 2.83e-12;  |   |      |          |
| Matches               | 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;                   |   |               | Matches               | 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;                   |   |      | Matches               | 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;                   |   |   |          | Matches               | 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;                   |   |      |          |
| Db                    | 552   | VPSYSFIR-AHDSEVQDLI                     | 569           | Qy                    | 1   | VPSYSFIRAHDSEVQDLI                      | 19   | RESULT                | 11  | PRELIMINARY;                            | PRT;                                    | 1590 AA. | RESULT                | 11  | PRELIMINARY;                            | PRT; | 1590 AA. |
| Qy                    |   |   |               | AC                    | 055263;   |   |      | AC                    | 055263;   |   |   |          | AC                    | 055263;   |   |      |          |
| DT                    | 01-NOV-1996   | (TREMBLREL. 01, CREATED)                |               | DT                    | 01-NOV-1996   | (TREMBLREL. 01, CREATED)                |      | DT                    | 01-NOV-1996   | (TREMBLREL. 01, CREATED)                |   |          | DT                    | 01-NOV-1996   | (TREMBLREL. 01, CREATED)                |      |          |
| DT                    | 01-JAN-1998   | (TREMBLREL. 05, LAST SEQUENCE UPDATE)   |               | DT                    | 01-JAN-1998   | (TREMBLREL. 05, LAST SEQUENCE UPDATE)   |      | DT                    | 01-JAN-1998   | (TREMBLREL. 05, LAST SEQUENCE UPDATE)   |   |          | DT                    | 01-JAN-1998   | (TREMBLREL. 05, LAST SEQUENCE UPDATE)   |      |          |
| DT                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |               | DT                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      | DT                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |   |          | DT                    | 01-NOV-1998   | (TREMBLREL. 08, LAST ANNOTATION UPDATE) |      |          |

DE GLUCOSYLTRANSFERASE GTF-I.  
 OS STREPTOCOCUS SORRINUS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.  
 [1] SEQUENCE FROM N.A.  
 RP STRAIN=ATCC 33478;  
 RA SATO S.; MONCHOIS V.; REMAUD-SIMEON M.; MONSAN P.; WILLEMET R.M.;  
 RL ANN. KAGOSHIMA UNIV. DENTAL SCHOOL 16:23-29(1996).  
 DR EMBL; D63570; D1010438; -.  
 DR PF00128; alpha-amylase; 1.  
 KW TRANSFERASE.  
 SQ SEQUENCE 1590 AA; 176057 MW; 47186FFA CRC32;  
 Query Match 71.9%; Score 97; DB 2; Length 1590;  
 Best Local Similarity 84.2%; Pred. No. 1.14e-08;  
 Matches 16; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
 QY 548 VPSYSFAR-AHSEVQDIT 565  
 1 VPSYSFIRTAHDSEVQDLI 19

RESULT 12 PRELIMINARY; PRT; 1590 AA.  
 ID 059983; ID 048756; PRELIMINARY; PRT; 1290 AA.  
 AC 059983; ID 048756; PRELIMINARY; PRT; 1290 AA.  
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
 DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)  
 DE GLUCOSYLTRANSFERASE-I PRECURSOR (EC 2.4.1.5) (DEXTRANSUCRASE) (SUCROSE  
 6-GLUCOSYLTRANSFERASE).  
 DE GTF-I.  
 OS STREPTOCOCCUS SORRINUS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAIN=OMZ176;  
 RX MEDLINE; 94146405.  
 RA SATO S.; INOUYE M.; HANADA N.; AIZAWA Y.; ISOBE Y.; KATAYAMA T.;  
 RT DNA sequence of the glucosyltransferase gene of serotype d  
 Streptococcus sobrinus.";  
 RT DNA SEQ. 4:19-27(1993).  
 CC D-FRUCTOSE + (1,6-ALPHA-D-GLUCOSYL)(N+1).  
 DR EMBL; D1003482; -.  
 DR PFAM; PF00128; alpha-amylase; 1.  
 DR SIGNAL; TRANSFERASE; GLYCOSYLTRANSFERASE.  
 FT SIGNAL 1 38 POTENTIAL.  
 SQ CHAIN 39 1590 AA; 175955 MW; F6E457D7 CRC32;

Query Match 71.9%; Score 97; DB 2; Length 1590;  
 Best Local Similarity 84.2%; Pred. No. 1.14e-08;  
 Matches 16; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

Db 548 VPSYSFAR-AHSEVQDIT 565  
 1 VPSYSFIRTAHDSEVQDLI 19

RESULT 13 PRELIMINARY; PRT; 1508 AA.  
 ID 052224; ID 052224; PRELIMINARY; PRT; 1449 AA.  
 AC 052224; ID 052224; PRELIMINARY; PRT; 1449 AA.  
 DT 01-JUN-1998 (TREMBLREL. 06, CREATED)  
 DT 01-JUN-1998 (TREMBLREL. 06, LAST SEQUENCE UPDATE)  
 DT 01-JUN-1998 (TREMBLREL. 06, LAST ANNOTATION UPDATE)  
 DE GLYCOSYLTRANSFERASE (EC 2.4.1.5) (DEXTRANSUCRASE)  
 DE (SUCROSE 6-GLYCOSYLTRANSFERASE).  
 GN DSRB.  
 OS LEUCONOSTOC MESENTEROIDES.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; LACTOBACILLACEAE;  
 OC LEUCONOSTOC.

RN [1] SEQUENCE FROM N.A.  
 RP STRAIN=NRRL B-1299;  
 RC MONCHOIS V.; REMAUD-SIMEON M.; MONSAN P.; WILLEMET R.M.;  
 RA FEMS MICROBIOL. LETT. 0:0-0(1998).  
 RL FEMS MICROBIOL. LETT. 0:0-0(1998).  
 CC -1- CATALYTIC ACTIVITY: SUCROSE +  
 CC (1,6-ALPHA-D-GLUCOSYL)(N) = D-FRUCTOSE +  
 CC (1,6-ALPHA-D-GLUCOSYL)(N+1).  
 DR EMBL; AF030129; G276612; -.  
 DR TRANSFERASE; GLYCOSYLTRANSFERASE.  
 KW SEQUENCE 1508 AA; 168511 MW; 7D3B6FFA CRC32;  
 SQ SEQUENCE 1508 AA; 168511 MW; 7D3B6FFA CRC32;  
 Query Match 69.6%; Score 94; DB 2; Length 1508;  
 Best Local Similarity 70.0%; Pred. No. 5.38e-03;  
 Matches 14; Conservative 4; Mismatches 1; Indels 1; Gaps 1;  
 Db 634 VPSYSFIR-AHSEVQVIA 652  
 QY 1 VPSYSFIRTAHDSEVQDLIA 20

RESULT 14 PRELIMINARY; PRT; 1290 AA.  
 ID 048756; ID 048756; PRELIMINARY; PRT; 1290 AA.  
 AC 048756; ID 048756; PRELIMINARY; PRT; 1290 AA.  
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE DEXTRANSUCRASE.  
 OS LEUCONOSTOC MESENTEROIDES.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; LACTOBACILLACEAE;  
 RN [1] LEUCONOSTOC.  
 RP STRAIN=NRRL B1299;  
 RX MONCHOIS V.; WILLEMET R.M.; REMAUD-SIMEON M.; CROUX C.; MONSAN P.;  
 RA CLoning and sequencing of a gene coding for a novel dextranucrase  
 from Leuconostoc mesenteroides NRRL B-1299 synthesizing only alpha  
 (1-6) and alpha (1-3) linkages. ;  
 RT GENE 112:23-32(1996).  
 RL EMBL; U38181; G1022963; -.  
 DR PFAM; PF00128; alpha-amylase; 1.  
 DR SEQUENCE 1290 AA; 145590 MW; F4460B13 CRC32;  
 SQ SEQUENCE 1290 AA; 145590 MW; F4460B13 CRC32;  
 Query Match 68.1%; Score 92; DB 2; Length 1290;  
 Best Local Similarity 78.9%; Pred. No. 1.99e-07;  
 Matches 15; Conservative 2; Mismatches 1; Indels 1; Gaps 1;  
 Db 388 VPSYSFIR-AHSEVQVIA 405  
 Qy 2 VPSYSFIRTAHDSEVQDLIA 20

RESULT 15 PRELIMINARY; PRT; 1449 AA.  
 ID 052264; ID 052264; PRELIMINARY; PRT; 1449 AA.  
 AC 052264; ID 052264; PRELIMINARY; PRT; 1449 AA.  
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE GLYCOSYLTRANSFERASE PRECURSOR.  
 GN GTF-I.  
 OS STREPTOCOCCUS SALIVARIUS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.  
 RN [1] SEQUENCE FROM N.A.  
 RP MEDLINE; 95122197.  
 RA SIMPSON C.L.; GIFFARD P.M.; JACQUES N.A.;  
 RT "Streptococcus salivarius ATCC 25975 possesses at least two genes  
 RT coding for primer-independent glucosyltransferases.";  
 RL INFECT. IMMUN. 63:609-621(1995).  
 DR EMBL; L35495; G662379; -.  
 DR PFAM; PF00128; alpha-amylase; 1.

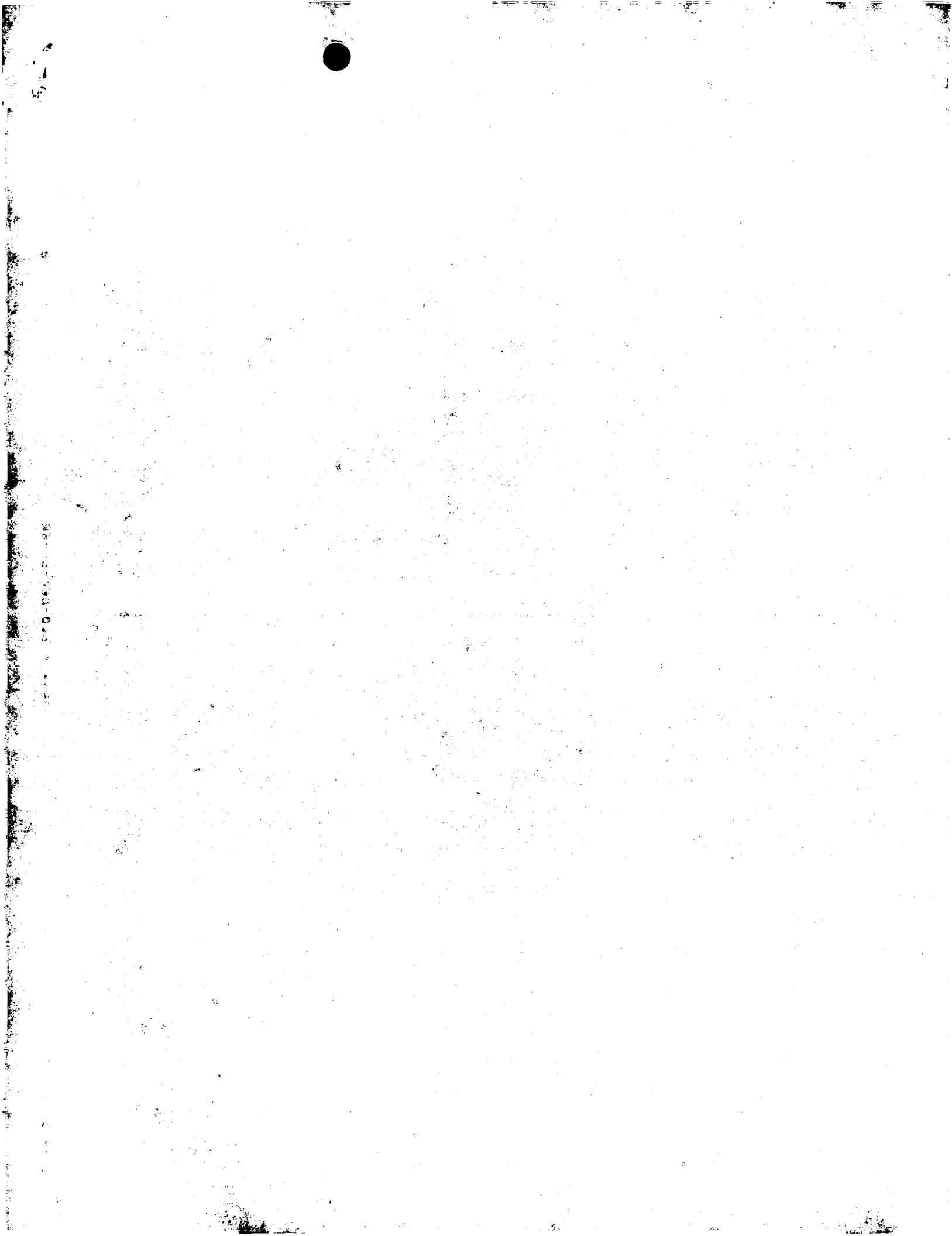
Thu Jan 13 09:13:40 2000

US-09-290-049-2.rspt

Page 5

KW SIGNAL; TRANSFERASE.  
FT SIGNAL 1 35 POTENTIAL.  
FT CHAIN 36 1449 GLUCOSYLTRANSFERASE.  
SQ SEQUENCE 1449 AA; 159984 MN; E9A4BA37 CRC32;  
Query Match 60.0%; Score 81; DB 2; Length 1449;  
Best Local Similarity 55.0%; Pred. No. 8 67e-05; 7; Mismatches 1; Indels 1; Caps 1;  
Matches 11; Conservative 1;  
Db 607 MANYAFVR-HDSEVQSIG 625  
QY ::|::|:|||||:|:  
1 VPSISFIRTAHDSEVQDLIA 20

Search completed: Tue Jan 11 15:45:45 2000  
Job time : 47 secs.





Query Match 100.0%; Score 155; DB 1; Length 1475;  
 Best Local Similarity 100.0%; Pred. No. 9.19e-23;  
 Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 1300 TGARTINGQOLYFRANGVQVG 1321  
 1 TGARTINGQOLYFRANGVQVG 22

RESULT 2  
 ID GTF1\_STRDO STANDARD; PRT; 1597 AA.  
 AC P11001;  
 DT 01-JUL-1989 (REL. 11, CREATED)  
 DT 01-JUL-1989 (REL. 11, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE GLUCOSYLTRANSFERASE-I PRECURSOR (EC 2.4.1.5) (GTF-I) (DEXTRANSUCRASE)  
 DE (SUCROSE 6-GLUCOSYLTRANSFERASE).  
 GN GTFI.  
 OS STREPTOCOCCUS DOWNEI (STREPTOCOCCUS SOBRINUS).  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.  
 [1]  
 RX MEDLINE; 87308014.  
 RA FERRETTI J.J., GILPIN M.L., RUSSELL R.R.B.;  
 RT "Nucleotide sequence of a glucosyltransferase gene from Streptococcus  
 sobrinus MRE28.";  
 RL 159:4271-4278(1987).  
 CC -I- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT  
 TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE  
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE  
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.  
 CC -I- CATALYTIC ACTIVITY: SUCROSE + (1,6-ALPHA-D-GLUCOSYL)(N) -  
 D-FRUCTOSE + (1,6-ALPHA-D-GLUCOSYL)(N+1).  
 CC -I- SUBCELLULAR LOCATION: SECRETED.  
 CC -I- DISEASE: DENTAL CARIES.  
 CC -I- GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED  
 GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES WATER-SOLUBLE  
 GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH FORMS OF  
 GLUCANS.  
 CC -I- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-  
 BINDING PROTEIN FROM S. MUTANS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).

CC EMBL; M17391; G153647; -  
 DR PFAM; PF00128; alpha-amylase; 1.  
 KW TRANSFERASE; GLYCOSYLTRANSFERASE; SIGNAL; REPEAT; DENTAL CARIES.  
 FT SIGNAL 1 38 POTENTIAL.  
 FT CHAIN 39 1597 GLYCOSYLTRANSFERASE-I.  
 FT DOMAIN 39 1050 CATALYTIC (APPROXIMATE).  
 FT DOMAIN 1099 1597 GLUCAN-BINDING (APPROXIMATE).  
 FT DOMAIN 1099 1597 1.25 A, 2 B, AND 5 AC REPEATS.  
 FT REPEAT 1099 1132 A REPEAT.

RESULT 3  
 ID GTF2\_STRDO STANDARD; PRT; 1592 AA.  
 AC P27470;  
 DT 01-AUG-1992 (REL. 23, CREATED)  
 DT 01-AUG-1992 (REL. 23, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE GLUCOSYLTRANSFERASE-I PRECURSOR (EC 2.4.1.5) (GTF-I) (DEXTRANSUCRASE)  
 DE (SUCROSE 6-GLUCOSYLTRANSFERASE).  
 OS STREPTOCOCCUS DOWNEI (STREPTOCOCCUS SOBRINUS).  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.  
 [1]  
 RN SEQUENCE FROM N. A.  
 RP STRAIN=6715;  
 RX MEDLINE; 91123227.  
 RA ABO H., MATSUMURA T., KODAMA T., OHTA H., FUKUI K., KATO K.,  
 RA KAGAWA H.;  
 RT "Peptide sequences for sucrose splitting and glucan binding within  
 Streptococcus sobrinus glucosyltransferase (water-insoluble glucan  
 synthetase);";  
 RL J. BACTERIOL. 173:989-996(1991).  
 CC -I- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT  
 TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE  
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE  
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.  
 CC -I- CATALYTIC ACTIVITY: SUCROSE + (1,6-ALPHA-D-GLUCOSYL)(N) -  
 D-FRUCTOSE + (1,6-ALPHA-D-GLUCOSYL)(N+1).  
 CC -I- SUBCELLULAR LOCATION: SECRETED.  
 CC -I- DISEASE: DENTAL CARIES.  
 CC -I- GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED  
 GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES WATER-SOLUBLE  
 GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH FORMS OF  
 GLUCANS.  
 CC -I- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-  
 BINDING PROTEIN FROM S. MUTANS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).

CC EMBL; D90213; G217033; -  
 DR PIR; A38175; A8815.  
 DR PFAM; PF00128; alpha.amylase; 1.  
 DR RSSP; P00655; 3HEE.  
 DR TRANSFERASE; GLYCOSYLTRANSFERASE; SIGNAL; REPEAT; DENTAL CARIES.  
 FT SIGNAL 1 38 POTENTIAL.  
 FT CHAIN 39 1592 GLYCOSYLTRANSFERASE-I.  
 FT DOMAIN 39 1044 CATALYTIC (APPROXIMATE).  
 FT DOMAIN 1093 1592 GLUCAN-BINDING (APPROXIMATE).  
 FT DOMAIN 1093 1592 6.5 X TANDEM REPEATS.

|        |   |  |                     |                  |
|--------|---|--|---------------------|------------------|
| FT     | SIGNAL  | 1  | 34                  |                  |
| FT     | REPEAT  | 1093   | 1142                | 1.               |
| FT     | REPEAT  | 1158   | 1207                | 2.               |
| FT     | REPEAT  | 1222   | 1272                | 3.               |
| FT     | REPEAT  | 1287   | 1337                | 4.               |
| FT     | REPEAT  | 1402   | 1451                | 5.               |
| FT     | REPEAT  | 1514   | 1563                | 6.               |
| FT     | REPEAT  | 1577   | 1592                | 7. (INCOMPLETE). |
| SO     | SEQUENCE  | 1592 AA;   | 176157 MW;          | 273A23FA CRC32;  |
| Db     | Query Match   | 72.3%  | Score 112; DB 1;    | Length 1592;     |
| Db     | Best Local Similarity   | 77.3%  | Pred. No. 1.49e-11; |                  |
| Db     | Matches   | 17;  | Conservative        |                  |
| Qy     | 1   | TGARTINGQOLYFRANGVQKG  | 22                  |                  |
| RESULT | 4   |  |                     |                  |
| ID     | GTFCSIRNU   | STANDARD;  | PRT;                | 1375 AA.         |
| AC     | P13470; P05427;   |  |                     |                  |
| DT     | 01-NOV-1988 (REL. 09, CREATED)  |  |                     |                  |
| DT     | 15-JAN-1990 (REL. 13, LAST SEQUENCE UPDATE)   |  |                     |                  |
| DE     | GLUCOSYLTRANSFERASE-SI PRECURSOR (EC 2.4.1.5) (GTF-SI)  |  |                     |                  |
| DE     | (DEXTRANSUCRASE) (SUCROSE 6-GLUCOSYLTRANSFERASE).   |  |                     |                  |
| GN     | GTFCS   |  |                     |                  |
| OS     | STREPTOCOCCUS MUTANS.   |  |                     |                  |
| OC     | BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCACEAE;  |  |                     |                  |
| OC     | STREPTOCOCCUS.  |  |                     |                  |
| RN     | [1]   |  |                     |                  |
| RP     | SEQUENCE FROM N.A.  |  |                     |                  |
| RC     | STRAININGS-5;   |  |                     |                  |
| RX     | MEDLINE; 89137980.  |  |                     |                  |
| RA     | UEDA S., SHIROZA T., KURAMITSU H. K.;   |  |                     |                  |
| RT     | "Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.";   |  |                     |                  |
| RL     | GENE 69:101-109(1988).  |  |                     |                  |
| RN     | [2]   |  |                     |                  |
| RP     | SEQUENCE OF 1-349 FROM N.A.   |  |                     |                  |
| RC     | STRAININGS-5;   |  |                     |                  |
| RX     | MEDLINE; 87308013.  |  |                     |                  |
| RA     | SHIROZA T., UEDA S., KURAMITSU H. K.;   |  |                     |                  |
| RT     | "Sequence analysis of the gtfB gene from Streptococcus mutans.;"  |  |                     |                  |
| CC     | J. BACTERIOL. 169:4263-4270(1987).  |  |                     |                  |
| CC     | - - FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE AGGRESSION OF BACTERIAL CELLS AND FOOD DEBRIS.   |  |                     |                  |
| CC     | INFECT. IMMUN. 58:2452-2458(1990).  |  |                     |                  |
| CC     | - - FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE AGGRESSION OF BACTERIAL CELLS AND FOOD DEBRIS.   |  |                     |                  |
| CC     | GILMORE K. S., RUSSELL R. R., FERRETTI J. J.:   |  |                     |                  |
| CC     | "Analysis of the Streptococcus downei gtfS gene, which specifies a glucosyltransferase that synthesizes soluble glucans.;"  |  |                     |                  |
| CC     | INFECT. IMMUN. 58:2452-2458(1990).  |  |                     |                  |
| CC     | - - FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE AGGRESSION OF BACTERIAL CELLS AND FOOD DEBRIS.   |  |                     |                  |
| CC     | D-FRUCTOSE + (1,6-ALPHA-D-GLUCOSYL)(N) = CATALYTIC ACTIVITY: SUCROSE + (1,6-ALPHA-D-GLUCOSYL) (N+1).  |  |                     |                  |
| CC     | - - CATALYTIC ACTIVITY: SUCROSE + (1,6-ALPHA-D-GLUCOSYL) (N) = D-FRUCTOSE + (1,6-ALPHA-D-GLUCOSYL) (N+1).   |  |                     |                  |
| CC     | - - ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF PRIMER GLUCAN UNLIKE GTF-I.  |  |                     |                  |
| CC     | - - DISEASE: DENTAL CARIES.   |  |                     |                  |
| CC     | - - SUBCELLULAR LOCATION: SECRETED.   |  |                     |                  |
| CC     | - - DISEASE: DENTAL CARIES.   |  |                     |                  |
| CC     | - - GLUCAN SYNTHETASE WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCANS AND SOME 1,6-LINKAGES), GTF-S SYNTHETIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHETIZES BOTH FORMS OF GLUCANS.  |  |                     |                  |
| CC     | - - SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.  |  |                     |                  |
| CC     | - - GLUCAN SYNTHETASE WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE).  |  |                     |                  |
| CC     | - - SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.  |  |                     |                  |
| CC     | This SWISS-PROT entry is copyright. It is produced through a collaboration between the European Bioinformatics Institute. There are no restrictions on its use by non profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch).   |  |                     |                  |
| CC     | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch). |  |                     |                  |
| CC     | DR  | EMBL; M30043; G153653; -.  |                     |                  |
| CC     | DR  | A41433; A41483;  |                     |                  |
| CC     | DR  | PFAM: PF00128; alpha-amylase; 1.                                 |                     |                  |
| CC     | DR  | TRANSFERASE; GLYCOSYLTRANSFERASE; SIGNAL; REPEAT; DENTAL CARIES. |                     |                  |
| CC     | FT  | CHAIN 37 1365  |                     |                  |
| CC     | FT  | DOMAIN 37 1050   |                     |                  |
| CC     | FT  | DOMAIN 1083 1365   |                     |                  |
| CC     | FT  | DOMAIN 1083 1365   |                     |                  |
| CC     | FT  | REPEAT 1083 1331   |                     |                  |
| FT     | REPEAT  | 1150 1199  | 2.                  |                  |
| FT     | REPEAT  | 1227 1330  | A REPET.            |                  |
| FT     | REPEAT  | 1303 1330  | AC REPET.           |                  |
| FT     | REPEAT  | 1375 AA; 153022 MW;  | 3EA372E CRC32;      |                  |
| Db     | Query Match   | 68.4%  | Score 106; DB 1;    | Length 1375;     |
| Db     | Best Local Similarity   | 68.2%  | Pred. No. 4.46e-10; |                  |
| Db     | Matches   | 15;  | Conservative        |                  |
| Qy     | 1   | TGARTINGQOLYFRANGVQKG  | 22                  |                  |
| RESULT | 5   |  |                     |                  |
| ID     | GTFSSIRDO   | STANDARD;  | PRT;                | 1365 AA.         |
| AC     | P29336;   |  |                     |                  |
| DT     | 01-DEC-1992 (REL. 24, CREATED)  |  |                     |                  |
| DT     | 01-DEC-1992 (REL. 24, LAST SEQUENCE UPDATE)   |  |                     |                  |
| DT     | 01-APR-1993 (REL. 25, LAST ANNOTATION UPDATE)   |  |                     |                  |
| DE     | GLUCOSYLTRANSFERASE-SI PRECURSOR (EC 2.4.1.5) (GTF-S) (DEXTRANSUCRASE).   |  |                     |                  |
| DE     | (SUCROSE 6-GLUCOSYLTRANSFERASE).  |  |                     |                  |
| DE     | (GTFCS).  |  |                     |                  |
| OS     | STREPTOCOCCUS DOWNEI (STREPTOCOCCUS SOBRINUS).  |  |                     |                  |
| OC     | BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCACEAE;  |  |                     |                  |
| OC     | STREPTOCOCUS.   |  |                     |                  |
| RN     | [1]   |  |                     |                  |
| RP     | SEQUENCE FROM N.A.  |  |                     |                  |
| RC     | STRAIN-MFP28;   |  |                     |                  |
| RX     | MEDLINE; 90316665.  |  |                     |                  |
| RA     | GILMORE K. S., RUSSELL R. R., FERRETTI J. J.:   |  |                     |                  |
| RT     | "Analysis of the Streptococcus downei gtfS gene, which specifies a glucosyltransferase that synthesizes soluble glucans.;"  |  |                     |                  |
| RT     | INFECT. IMMUN. 58:2452-2458(1990).  |  |                     |                  |
| CC     | - - FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE AGGRESSION OF BACTERIAL CELLS AND FOOD DEBRIS.   |  |                     |                  |
| CC     | D-FRUCTOSE + (1,6-ALPHA-D-GLUCOSYL)(N) = CATALYTIC ACTIVITY: SUCROSE + (1,6-ALPHA-D-GLUCOSYL) (N+1).  |  |                     |                  |
| CC     | - - ENZYME REGULATION: GLUCAN SYNTHESIS BY GTF-S IS INDEPENDENT OF PRIMER GLUCAN UNLIKE GTF-I.  |  |                     |                  |
| CC     | - - DISEASE: DENTAL CARIES.   |  |                     |                  |
| CC     | - - GLUCAN SYNTHETASE WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE).  |  |                     |                  |
| CC     | - - SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.  |  |                     |                  |
| CC     | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch). |  |                     |                  |
| CC     | DR  | M30043; G153653; -.  |                     |                  |
| CC     | DR  | A41433; A41483;  |                     |                  |
| CC     | DR  | PFAM: PF00128; alpha-amylase; 1.                                 |                     |                  |
| CC     | DR  | TRANSFERASE; GLYCOSYLTRANSFERASE; SIGNAL; REPEAT; DENTAL CARIES. |                     |                  |
| FT     | CHAIN 35 1050   |  |                     |                  |
| FT     | DOMAIN 1126 1375  |  |                     |                  |
| FT     | DOMAIN 1126 1375  |  |                     |                  |
| FT     | REPEAT 1126 1159  |  |                     |                  |
| FT     | REPEAT 1127 1330  |  |                     |                  |
| FT     | REPEAT 1227 1330  |  |                     |                  |
| FT     | REPEAT 1303 1330  |  |                     |                  |
| FT     | REPEAT 1375 AA; 153022 MW;  |  |                     |                  |
| Db     | Query Match   | 68.4%  | Score 106; DB 1;    | Length 1375;     |
| Db     | Best Local Similarity   | 68.2%  | Pred. No. 4.46e-10; |                  |
| Db     | Matches   | 15;  | Conservative        |                  |
| Qy     | 1   | TGARTINGQOLYFRANGVQKG  | 22                  |                  |

FT REPEAT 1225 1274 3.  
 FT REPEAT 1289 1339 4.  
 FT REPEAT 1353 1365 5 (INCOMPLETE).  
 SQ SEQUENCE 1365 AA; 151590 MW; 156FB8B31 CRC32;

Query Match Similarity 63.6%; Score 104; DB 1; Length 1365;  
 Best Local Similarity 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
 Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 1235 TGEQTIDGQKVFFQDNGVQVK 1256  
 QY 1 TGARTINGQOLLYFRANGVQVK 22

RESULT 6  
 ID GTFD\_STRMU STANDARD; PRT; 1430 AA.  
 AC P49331;  
 DT 01-FEB-1995 (REL. 33, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 RA KATSOULOU C., TZERIMIA M., TAVERNARAKIS N., ALEXANDRAKI D.;  
 DE GLUCOSYLTRANSFERASE-S PRECURSOR (EC 2.4.1.5) (GTF-S) (DEXTRANSUCRASE)  
 DE (SUCROSE 6-GLUCOSYLTRANSFERASE).  
 RT GTFD.  
 OS STREPTOCOCCUS MITANS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAININGS-5;  
 RX MEDLINE; 91100958.  
 RA HONDA O., KATO C., KURAMITSU H.K.;  
 RT J. GEN. MICROBIOL. 136:2090-2105(1990).  
 CC -I- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT  
 TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE  
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE  
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.  
 CC -I- CATALYTIC ACTIVITY: SUCROSE + (1,6-ALPHA-D-GLUCOSYL)(N) -  
 CC D-FRUCTOSE + (1,6-ALPHA-D-GLUCOSYL)(N+1).  
 CC -I- SUBCELLULAR LOCATION: SECRETED.  
 CC -I- DISEASE: DENTAL CARIES.  
 CC -I- GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED  
 GLUCOSE AND SOME 1,6 LINKAGES); GTF-S SYNTHESIZES WATER-SOLUBLE  
 GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH FORMS OF  
 GLUCANS.  
 CC -I- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-  
 CC -I- BINDING PROTEIN FROM S. MUTANS.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use. By non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial/  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).

RESULT 7  
 ID YJ07\_FEAST STANDARD; PRT; 382 AA.  
 AC P47007;  
 DT 01-FEB-1995 (REL. 33, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE HYPOTHETICAL 44.9 KD PROTEIN IN INO1-IDS2 INTERGENIC REGION.  
 GN YJ114C OR J0639.  
 OS SACCHAROMYCETES CEREVISIAE (BAKER'S YEAST).  
 OC EUKARYOTA; FUNGI; ASCOMYCOTA; HEMIASCOMYCETES; SACCHAROMYCETALES;  
 OC SACCHAROMYCETACEA; SACCHAROMYCES.  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAIN-S288C / FYI1679;  
 RX MEDLINE; 98408771.  
 RA KATSOULOU C., TZERIMIA M., TAVERNARAKIS N., ALEXANDRAKI D.;  
 RT "Sequence analysis of a 40.7 kb segment from the left arm of yeast  
 chromosome X reveals 14 known genes and 13 new open reading frames  
 including homologues of genes clustered on the right arm of  
 chromosome XI.";  
 RT YEAST 12:787-791(1998).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use. By non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial/  
 entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).  
 DR EMBL; X87371; G884550; -.  
 DR EMBL; 249422; G101561; -.  
 SQ HYPOTHETICAL PROTEIN.  
 KW SEQUENCE 382 AA; 44862 MW; 5B128AB9 CRC32;

Query Match Similarity 51.0%; Score 79; DB 1; Length 382;  
 Best Local Similarity 38.9%; Pred. No. 7.50e-04; Indels 0; Gaps 0;  
 Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;  
 DE DEATH ON CURING PROTEIN.  
 GN DOC.  
 OS BACTERIOPHAGE PI.  
 OC VIRUSES; DS DNA VIRUSES, NO RNA STAGE; TAILED PHAGES; MYOVIRIDAE.  
 RN [1] SEQUENCE FROM N.A.  
 RP MEDLINE; 94016501.  
 RX LEHNHEER H., MAGUIN E., JAFRI S., YARMOLINSKY M.B.;  
 RT "Plasmid addiction genes of bacteriophage PI: doc, which causes cell  
 death on curing of prophage, and phd, which prevents host death when  
 prophage is retained";  
 RT J. MOL. BIOL. 233:414-428(1993).  
 RL -I- FUNCTION: KILLS CELLS. DOC AND PHD PROTEIN FUNCTION IN UNISON TO  
 CC -I- STABILIZE PLASMID NUMBER BY INDUCING A LETHAL RESPONSE TO PLASMID  
 CC LOSS.  
 CC -I- THE CONCENTRATION OF PHD IN PI LYSOGENS IS FAR GREATER THAN THAT  
 CC OF THE POISON IT ANTAGONIZES. SUCH AN EXCESS MAY ASSUME THE  
 CC WELL-BEING OF CARRIERS OF THE ADDICTING PLASMID.

Query Match Similarity 60.0%; Score 93; DB 1; Length 1430;  
 Best Local Similarity 59.1%; Pred. No. 5.53e-07; Indels 0; Gaps 0;  
 Matches 13; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 1332 TGQTTAGKKLYFASDQKVKG 1353  
 CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: M95666; G463277; -.

CC PIR: S38555; S38555;

DR SEQUENCE 126 AA; 13588 MW; E4600881 CRC32;

DR QY 5 TINGOLLYFRANGVQV 20

Query Match 46.5%; Score 72; DB 1; Length 126; Best Local Similarity 56.3%; Pred. No. 2.20e-02; Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Db 76 ALNSALFLRNGVQV 91

F T 9  
I Y119 NPVOP STANDARD; PRT; 529 AA.  
AC 010358;  
DT 01-NOV-1997 (REL. 35, CREATED)  
DT 01-NOV-1997 (REL. 35, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

DE HYPOTHETICAL 59.0 KD PROTEIN PRECURSOR (ORF119).

OS ORGYIA PSEUDOTSGATA MULTICAPSID POLYBHDROSIS VIRUS (OPMNPV).

OC VIRUSES; DSDNA VIRUSES, NO RNA STAGE; BACULOVIRIDAE;

RN [1]

RP SEQUENCE FROM N.A.

RA MEDLINE: 90129305.

RX DOVE C.H., WANG S.Z., PRICE S.B., PHELPHS C.J., LYERLY D.M.,  
RA WILKINS T.W., JOHNSON J.L.;  
RT "Molecular characterization of the Clostridium difficile toxin A gene.";  
RT INPECT. IMMUN. 58:480-488(1990).

RN [33]

RP SEQUENCE FROM N.A.

RC STRAINNPVI 10463;

RA VON EICHEL-STREIBER C;

RL SUBMITTED (JAN-1997) TO EMBL/GENBANK/DDJB DATA BANKS.

CC -1- DOMAIN: THE C-TERMINAL PART OF TOXIN A IS COMPOSED OF FIVE REPETITIVE STRUCTURE. THIS PART OF TOXIN A IS COMPOSED OF FIVE DIFFERENT OLIGOPEPTIDES.

CC -1- DISEASE: CLOSTRIDIUM DIFFICILE PRODUCES TWO TOXINS, AN ENTEROTOXIN CALLED A AND CYTOTOXIN B. ONLY AFTER THE ENTERAL DELIVERY OF THE ENTEROTOXIN MAY THE CHARACTERISTIC DISEASE CALLED PSEUDOMEMBRANOUS COLITIS BE INDUCED.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: X51797; G40441; -.

DR EMBL: M30307; G144926; -.

DR PIR: S0863B; S0863B.

RW TOXIN; ENTEROTOXIN.

SQ SEQUENCE 2710 AA; 308052 MW; D04C4C08 CRC32;

Query Match 43.2%; Score 67; DB 1; Length 2710; Best Local Similarity 52.2%; Pred. No. 2.19e-01; Matches 12; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

Db 2460 SGRRTSGLRFHADGEVSG 280

F T CHAIN 21 529 HYPOTHETICAL PROTEIN ORF119.

SQ SEQUENCE 529 AA; 58983 MW; B48905EA CRC32;

Query Match 45.2%; Score 70; DB 1; Length 529; Best Local Similarity 42.9%; Pred. No. 5.58e-02; Matches 9; Conservative 7; Mismatches 5; Indels 0; Gaps 0;

Db 260 SGRRTSGLRFHADGEVSG 280

F T CHAIN 21 529 HYPOTHETICAL PROTEIN ORF119.

SQ SEQUENCE 529 AA; 58983 MW; B48905EA CRC32;

RESULT 10  
ID TOXA\_CLODI STANDARD; PRT; 2710 AA.  
AC P16154;  
DT 01-APR-1990 (REL. 14, CREATED)  
DT 01-EB-1995 (REL. 33, LAST SEQUENCE UPDATE)  
DT 01-NOV-1997 (REL. 35, LAST ANNOTATION UPDATE)

Db 260 SGRRTSGLRFHADGEVSG 280

F T CHAIN 21 529 HYPOTHETICAL PROTEIN ORF119.

SQ SEQUENCE 529 AA; 58983 MW; B48905EA CRC32;

RESULT 11  
ID PTER\_MOUSE STANDARD; PRT; 349 AA.  
AC 060866;  
DT 15-DEC-1998 (REL. 37, CREATED)  
DT 15-DEC-1998 (REL. 37, LAST SEQUENCE UPDATE)  
DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)

DE PHOSPHOTRIESTERASE RELATED PROTEIN (PARATHION HYDROLASE-RELATED PROTEIN).  
DE PTER.

GN MUS MUSCULUS (MOUSE).  
OS EUKARYOTA; METAEOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
RN [1]  
RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE-KIDNEY;

RX MEDLINE: 96194899.

RA HOU X., MASER R.L., MAGENHEIMER B.S., CALVET J.P.;  
RT "A mouse kidney- and liver-expressed cDNA having homology with a CLOSTRIDIUM DIFFICILE."

RT prokaryotic partition hydrolase (phosphotriesterase)-encoding gene;  
 RT abnormal expression in injured and polycystic kidneys.";  
 RL EMBL; X69460; G15415; -.  
 CC EMBL; AJ001191; E1231676; -.  
 CC -I- COPACTOR: CONTAINS 2 MOLES OF ZINC PER SUBUNIT (BY SIMILARITY).  
 CC -I- TISSUE SPECIFICITY: EXPRESSED IN KIDNEY AND LIVER.  
 CC -I- SIMILARITY: BELONGS TO THE PHOSPHOESTERASE FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; U28165; G881499; -.  
 DR MGD; MGI:107372; PTER.  
 KW HYDROLASE; ZINC.  
 METAL 26 26 ZINC 1 (BY SIMILARITY).  
 METAL 28 28 ZINC 1 (BY SIMILARITY).  
 METAL 169 169 ZINC 1 AND 2 (BY SIMILARITY).  
 METAL 201 201 ZINC 2 (BY SIMILARITY).  
 METAL 230 230 ZINC 2 (BY SIMILARITY).  
 METAL 298 298 ZINC 1 (BY SIMILARITY).  
 SEQUENCE 349 AA: 39218 MW: 58404 EEF CRC32;

Query Match 41.9%; Score 65; DB 1; Length 1396;  
 Best Local Similarity 53.8%; Pred. No. 5.32e-01;  
 Matches 7; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 79 AIREELVFKAG 91  
 Qy :|:||||:|| 5 TINGQQLYFRANG 17

RESULT 12 STANDARD; PRT; 1396 AA.  
 ID VLT\_PTS  
 AC P13390; O48502;  
 DT 01-JAN-1990 (REL. 13, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST SEQUENCE UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE L-SHAPED TAIL FIBRE PROTEIN (LTF PROTEIN).  
 OS BACTERIOPHAGE T5.  
 OC VIRUSES; DSrna VIRUSES, NO RNA STAGE; TAILED PHAGES; SIPHOVIRIDAE.  
 RN [1] SEQUENCE FROM N.A.  
 RP MEDLINE; 95309401.  
 RA KALIMAN A.V., KULSHIN V.E., SHLYAKHNIKOV M.G., KSENZENKO V.N.,  
 RA KRYUKOV V.M.;  
 RA "The nucleotide sequence of the bacteriophage T5 ltf gene.";  
 RL FEBS LETT. 366:46-48(1995).  
 RN SEQUENCE FROM N.A.  
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.  
 RA KALIMAN A.V., KRYUKOV V.M., BAYEV A.A.;  
 RL SUBMITTED (JAN-1998) TO EMBL/GENBANK/DDJB DATA BANKS.  
 RN [3] MEDLINE; 88389370.  
 RP PRELIMINARY PARTIAL SEQUENCE FROM N.A.  
 RA KALIMAN A.V., KRYUKOV V.M., BAYEV A.A.;  
 RL "The nucleotide sequence of bacteriophage T5 DNA at the region between early and late genes.";  
 RL NUCLEIC ACIDS RES. 16:6230-6230(1988).  
 CC -I- FUNCTION: NONSPECIFIC PROTEIN THAT MEDIATES BINDING TO THE POLYMANNOSE O ANTIGEN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

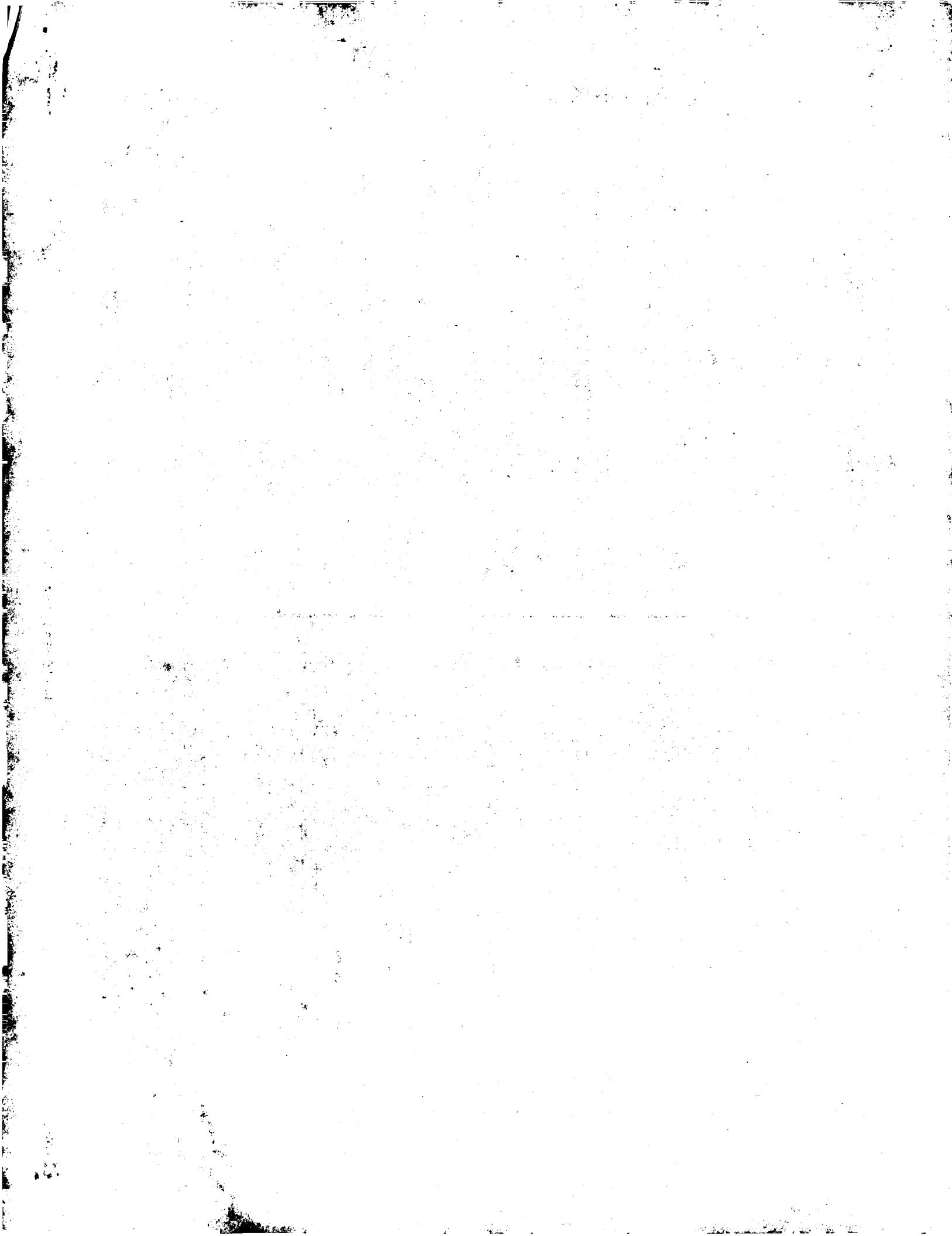
DR EMBL; X53042; G53105; -.  
 DR PIR; S13155; S13165.  
 DR MGD; MGI:88082; ASGR2.  
 DR PROST1; PS00615; C\_TYPELECTIN\_1; 1.  
 DR PROST1; PS50041; C\_TYPELECTIN\_2; 1.  
 DR PFAM; PF00059; lectin\_c; 1.  
 DR HSSP; P0674; IRIKE.  
 DR LECTIN; GLYCOPROTEIN; RECEPTOR; ENDOCYTOSIS; TRANSMEMBRANE;  
 KW CALCIUM; SIGNAL-ANCHOR; PHOSPHORYLATION.  
 FT DOMAIN 1 58 CYTOPLASMIC (POTENTIAL).  
 FT TRANSM 59 79 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).  
 FT DOMAIN 80 301 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 169 295 C-TYPE LECTIN (LONG FORM).  
 FT DISULFID 170 181 BY SIMILARITY.  
 FT DISULFID 198 293 BY SIMILARITY.

CC RESULT 13 STANDARD; PRT; 301 AA.  
 ID LEC1\_MOUSE  
 AC P24721;  
 DT 01-MAR-1992 (REL. 21, CREATED)  
 DT 01-MAR-1992 (REL. 21, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)  
 DE ASIALOGLYCOPEPTIDE RECEPTOR 2 (HEPATIC LECTIN 2) (ASGPR).  
 OS MUS MUSCULUS (MOUSE).  
 OC EURARCOTA; METAORA; CHORDATA; VERTEBRATA; MAMMALIA; EUOTHERIA;  
 OC RODENTIA; SCIUROGNATHI; MURIDAE; MURINAE; MUS.  
 RN [1] SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6; TISSUE=LIVER;  
 RX MEDLINE; 91027942.  
 RA SANFORD J.P., DOYLE D.;  
 RT "Mouse asialoglycoprotein receptor cDNA sequence: conservation of receptor genes during mammalian evolution.";  
 RL BIOMIM. BIOPHYS. ACTA 1087:259-261(1990).  
 CC WHICH THE TERMINAL STYLIC ACID RESIDUE ON THEIR COMPLEX CARBOHYDRATE MOieties HAS BEEN REMOVED. THE RECEPTOR RECOGNIZES TERMINAL GALACTOSE AND N-ACTYLGALACTOSAMINE UNITS. AFTER LIGAND BINDING TO THE RECEPTOR, THE RESULTING COMPLEX IS INTERNALIZED AND TRANSPORTED TO A SORTING ORGANELLE, WHERE RECEPTOR AND LIGAND ARE DISASSOCIATED. THE RECEPTOR THEN RETURNS TO THE CELL MEMBRANE SURFACE.  
 CC -I- SUBCELLULAR LOCATION: TYPE II MEMBRANE PROTEIN.  
 CC -I- TISSUE SPECIFICITY: EXPRESSED EXCLUSIVELY IN HEPATIC PARENCHYMAL CELLS.  
 CC -I- CALCIUM IS REQUIRED FOR LIGAND BINDING.  
 CC -I- SIMILARITY: CONTAINS 1 C-TYPE LECTIN FAMILY DOMAIN.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; X53042; G53105; -.  
 DR PIR; S13155; S13165.  
 DR MGD; MGI:88082; ASGR2.  
 DR PROST1; PS00615; C\_TYPELECTIN\_1; 1.  
 DR PROST1; PS50041; C\_TYPELECTIN\_2; 1.  
 DR PFAM; PF00059; lectin\_c; 1.  
 DR HSSP; P0674; IRIKE.  
 DR LECTIN; GLYCOPROTEIN; RECEPTOR; ENDOCYTOSIS; TRANSMEMBRANE;  
 KW CALCIUM; SIGNAL-ANCHOR; PHOSPHORYLATION.  
 FT DOMAIN 1 58 CYTOPLASMIC (POTENTIAL).  
 FT TRANSM 59 79 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN).  
 FT DOMAIN 80 301 EXTRACELLULAR (POTENTIAL).  
 FT DOMAIN 169 295 C-TYPE LECTIN (LONG FORM).  
 FT DISULFID 170 181 BY SIMILARITY.  
 FT DISULFID 198 293 BY SIMILARITY.







Query Match 92.3%; Score 143; DB 2; Length 1476;  
 Best Local Similarity 95.5%; Pred. No. 5.35e-19; 0; Mismatches 1; Indels 0; Gaps 0;  
 Matches 21; Conservative

Db 1301 TGARTINGQHLYFRANGQVKG 1322  
 Qy 1 TGARTINGQHLYFRANGQVKG 22

RESULT 3 PRELIMINARY; PRT; 1476 AA.

ID 069381 AC 069381; DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE GLUCOSYLTRANSFERASE-I.  
 GN GTFB.  
 OS STREPTOCOCCUS MUTANS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA KIMURA S., HAMADA S.;  
 RT "Molecular analyses of glucosyltransferase genes among strains of  
 Streptococcus mutans";  
 RL FEMS MICROBIOL. LETT. 161:331-336(1998).  
 DR EMBL; D89651; D1027041; -.

RC STRAIN-MT8148;  
 RX FUTIWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S.,  
 RA KIMURA S., HAMADA S.;  
 RT "Molecular analyses of glucosyltransferase genes among strains of  
 Streptococcus mutans";  
 RL FEMS MICROBIOL. LETT. 161:331-336(1998).  
 DR EMBL; D89651; D1027041; -.

KW TRANSFERASE.

SQ SEQUENCE 1476 AA; 165638 MW; 0868C20A CRC32;

Query Match 92.3%; Score 143; DB 2; Length 1476;  
 Best Local Similarity 95.5%; Pred. No. 5.35e-19; 0; Mismatches 1; Indels 0; Gaps 0;  
 Matches 21; Conservative

Db 1301 TGARTINGQHLYFRANGQVKG 1322  
 Qy 1 TGARTINGQHLYFRANGQVKG 22

RESULT 4 PRELIMINARY; PRT; 1476 AA.

ID 069387 AC 069387; DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE GLUCOSYLTRANSFERASE-I.  
 GN GTFB.  
 OS STREPTOCOCCUS MUTANS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.

RN [1]  
 RP SEQUENCE FROM N.A.

RA KIMURA S., HAMADA S.;  
 RT "Molecular analyses of glucosyltransferase genes among strains of  
 Streptococcus mutans";  
 RL FEMS MICROBIOL. LETT. 161:331-336(1998).  
 DR EMBL; D89658; D1027050; -.

KW TRANSFERASE.

SQ SEQUENCE 1476 AA; 165647 MW; 258D4ADD CRC32;

Query Match 92.3%; Score 143; DB 2; Length 1476;  
 Best Local Similarity 95.5%; Pred. No. 5.35e-19; 0; Mismatches 1; Indels 0; Gaps 0;  
 Matches 21; Conservative

Db 1301 TGARTINGQHLYFRANGQVKG 1322  
 Qy 1 TGARTINGQHLYFRANGQVKG 22

RESULT 5 PRELIMINARY; PRT; 1476 AA.

ID 069390 AC 069390; DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE GLUCOSYLTRANSFERASE-I.  
 GN GTFB.  
 OS STREPTOCOCCUS MUTANS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.

RA KIMURA S., HAMADA S.;  
 RT "Molecular analyses of glucosyltransferase genes among strains of  
 Streptococcus mutans";  
 RL FEMS MICROBIOL. LETT. 161:331-336(1998).  
 DR EMBL; D88660; D1027053; -.

KW TRANSFERASE.

SQ SEQUENCE 1476 AA; 165633 MW; 9119C72B CRC32;

Query Match 92.3%; Score 143; DB 2; Length 1476;  
 Best Local Similarity 95.5%; Pred. No. 5.35e-19; 0; Mismatches 1; Indels 0; Gaps 0;  
 Matches 21; Conservative

Db 1301 TGARTINGQHLYFRANGQVKG 1322  
 Qy 1 TGARTINGQHLYFRANGQVKG 22

RESULT 6 PRELIMINARY; PRT; 1455 AA.

ID 069388 AC 069388; DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

DE GLUCOSYLTRANSFERASE-SI.  
 GN GTFC.  
 OS STREPTOCOCCUS MUTANS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.

RN [1]  
 RP SEQUENCE FROM N.A.

RA KIMURA S., HAMADA S.;  
 RT "Molecular analyses of glucosyltransferase genes among strains of  
 Streptococcus mutans";  
 RL FEMS MICROBIOL. LETT. 161:331-336(1998).  
 DR EMBL; D88658; D1027050; -.

KW TRANSFERASE.

SQ SEQUENCE 1455 AA; 163046 MW; 7659CFBC CRC32;

Query Match 78.1%; Score 121; DB 2; Length 1455;



Query Match 70.3%; Score 109; DB 2; Length 1518;  
Best Local Similarity 72.7%; Pred. No. 2.00e-10; Mismatches 4; Indels 0; Gaps 0;

Db 1383 TGAQVINGQHLYFNADGSQVKG 1404  
 1 TGARTINGQOLIYFRANGVQKG 22

RESULT 12 PRELIMINARY; PRT; 1518 AA.

ID Q05263 PRELIMINARY; PRT; 1590 AA.

AC Q05263; OC

DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
01-JAN-1998 (TREMBLREL. 05, LAST SEQUENCE UPDATE)

DE GLUCOSYLTRANSFERASE I (EC 2.4.1.5) (GTF) (DEXTRANSGUARASE).

GN (SUCROSE 6-GLUCOSYLTRANSFERASE).

OS STREPTOCOCCUS SALIVARIUS.

RA GIFFARD P.M.; SIMPSON C.L.; MILWARD C.P.; JACQUES N.A.;  
BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;

RT STREPTOCOCCUS.  
[1] SEQUENCE FROM N.A.

RQ STRAIN=ATCC 25975;

MEDLINE; 92148377;

RA GIFFARD P.M.; SIMPSON C.L.; MILWARD C.P.; JACQUES N.A.;  
BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;

RT glucoamyltransferase genes in *Streptococcus salivarius* ATCC 25975.;"

J. GEN. MICROBIOL. 137:2577-2593(1991).

"Molecular characterization of a cluster of at least two genes involved in the production of extracellular glucans, that are thought to play a key role in the development of the dental plaque because of their ability to adhere to smooth surfaces and mediate the aggregation of bacterial cells and food debris.

-1 CATALYTIC ACTIVITY: SUCROSE + (1,6-ALPHA-D-GLUCOSYL)(N) - D-FRUCTOSE + (1,6-ALPHA-D-GLUCOSYL)(N+1).  
-1 SUBCELLULAR LOCATION: EXTRACELLULAR.  
-1 DISEASE: DENTAL CARIES.

EMBL: Z11873; GI: 67527;  
EMBL: M64111; GI: 53649; -

DR PFAM: PF00118; alpha-amylase; 1.  
KW TRANSFERASE; GLYCOSYLTRANSFERASE; REPEAT; DENTAL CARIES.

FT DOMAIN 1307 1482 6 DIRECT REPEATS.

FT REPEAT 1307 1338 PEPT1.

FT REPEAT 1339 1352 PEPT2.

FT REPEAT 1372 1403 PEPT3.

FT REPEAT 1404 1417 PEPT4.

FT REPEAT 1437 1468 PEPT5.

FT REPEAT 1469 1482 PEPT6.

FT SEQUENCE 1518 AA; 167730 MW; 20728800 CRC32;

Query Match 70.3%; Score 109; DB 2; Length 1518;  
Best Local Similarity 72.7%; Pred. No. 2.00e-10; Mismatches 4; Indels 0; Gaps 0;

Db 1383 TGAQVINGQHLYFNADGSQVKG 1404  
 1 TGARTINGQOLIYFRANGVQKG 22

RESULT 13 PRELIMINARY; PRT; 1290 AA.

ID Q48756 PRELIMINARY; PRT; 1290 AA.

AC Q48756; OC

DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)

DE 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)

RQ DE DEXTRANSGUARASE.

RA LEUCONOSTOC MESENTEROIDES.

RA BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; LACTOBACILLACEAE;

RT LEUCONOSTOC.

RQ SEQUENCE FROM N.A.

RA STRAIN=NRL B1299;

MEDLINE; 9713666.

RA MONCHIOIS V.; WILLEMET R.M.; REINAUD-SIMEON M.; CROUX C.; MONSAN P.;  
"Cloning and sequencing of a gene coding for a novel dextransguarase from Leuconostoc mesenteroides NRRL B-1299 synthesizing only alpha (1-6) and alpha (1-3) linkages";

DR GENE 182:23-32(1996).

DR EMBL: U38181; GI: 6022963; -.

DR PFAM: PF00128; alpha-amylase; 1.

DR SEQUENCE 1290 AA; 145590 MW; F4460B13 CRC32;

Query Match 69.0%; Score 107; DB 2; Length 1290;  
Best Local Similarity 63.6%; Pred. No. 6.24e-10; Mismatches 5; Indels 0; Gaps 0;

Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Db 1181 TGDRTIKGOKLXFNONGIOMKG 1202  
 QY 1 TGARTINGQOLIYFRANGVQKG 22

RESULT 14 PRELIMINARY; PRT; 1390 AA.

ID 069385 PRELIMINARY; PRT; 1390 AA.

AC 069385; OC

DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
01-NOV-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)

DE GLUCOSYLTRANSFERASE-ST.

GN GTFc.

RA STREPTOCOCCUS MITANS.

RA BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;

RT STREPTOCOCCUS MITANS.

RQ SEQUENCE FROM N.A.

RA STRAIN=ATCC 33478;

RA SATO S.; ANN KAGOSHIMA UNIV. DENTAL SCHOOL 16:23-29(1996).

RA DR EMBL: D63570; D1010438; -.

DR PFAM: PF00128; alpha-amylase; 1.

DR SEQUENCE 1590 AA; 176057 MW; 47186F6A CRC32;

Query Match 70.3%; Score 109; DB 2; Length 1590;  
Best Local Similarity 68.2%; Pred. No. 2.05e-10; Mismatches 4; Indels 0; Gaps 0;

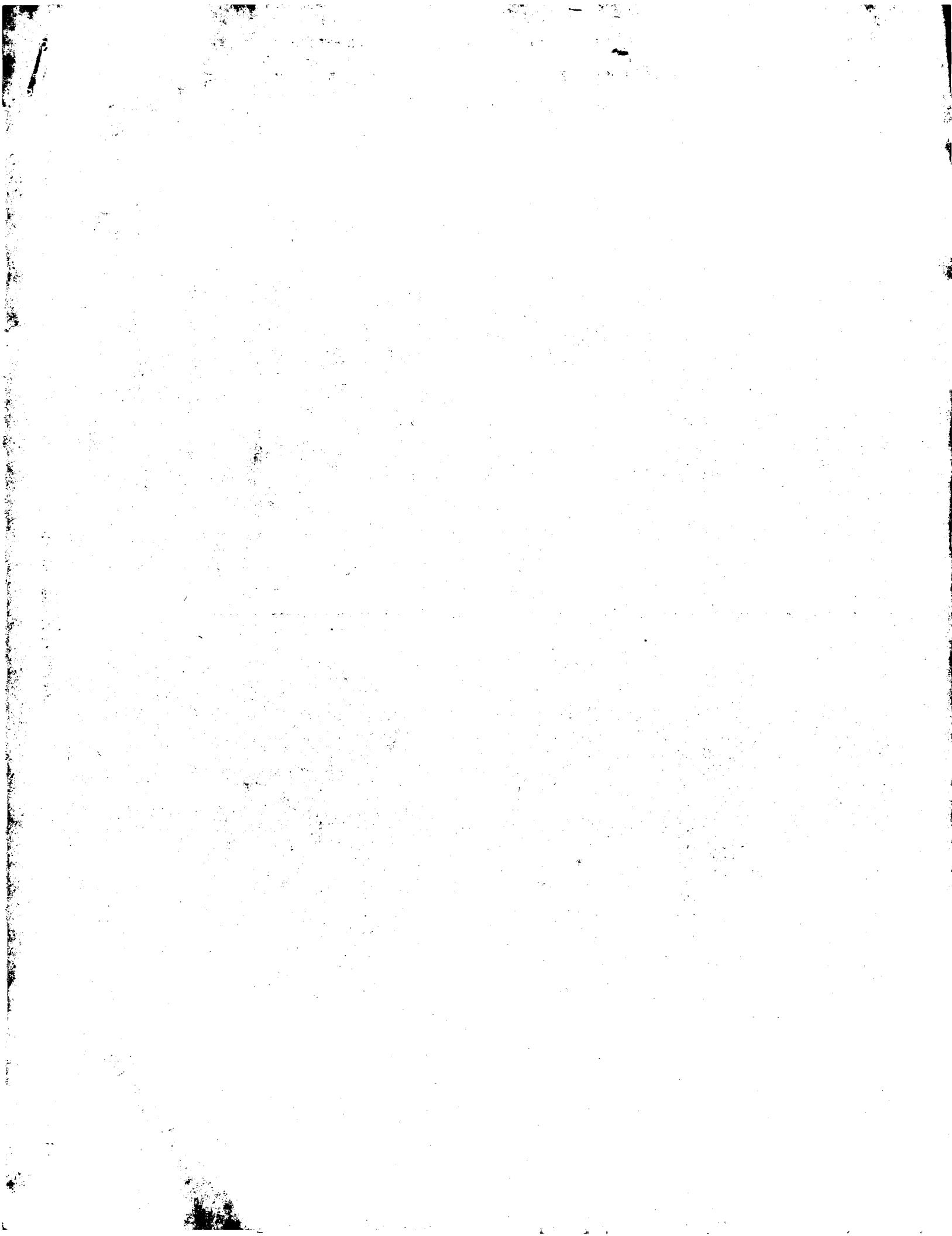
Db 1264 TGAQVINGQHLYFNADGSQVKG 1285  
 QY 1 TGARTINGQOLIYFRANGVQKG 22

RESULT 15 PRELIMINARY; PRT; 1599 AA.

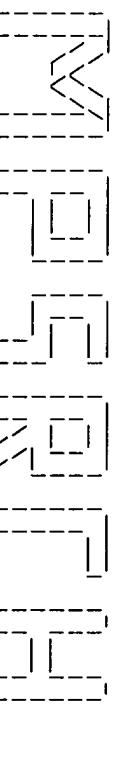
ID Q00599 PRELIMINARY; PRT; 1599 AA.

AC 000599;  
 DT 01-NOV-1996 (TREMBREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMBREL. 01, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBREL. 08, LAST ANNOTATION UPDATE)  
 DE GLUCOSYLTRANSFERASE S PRECURSOR (EC 2.4.1.5) (GTF) (DEXTRANSUCRASE)  
 DE (SUCROSE 6-GLUCOSYLTRANSFERASE).  
 GN GTFK.  
 OS STREPTOCOCCUS SALIVARIUS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.  
 RN [1]  
 RP STRAIN FROM N.A.  
 RC STRAIN=ATCC 25975;  
 RC MEDLINE; 93381463.  
 RA GIFFARD P.M., SIMPSON C.L., MILFARD C.P., SIMPSON C.L., JACQUES N.A.;  
 RT Molecular characterization of a cluster of at least two  
 RT glucosyltransferase genes in *streptococcus salivarius* ATCC 25975.;  
 RL J. GEN. MICROBIOL. 137:257-259(1991).  
 CC "Sequence of the gtf gene of *streptococcus salivarius*. and  
 evolution of the gtf genes of oral streptococci.";  
 RT J. GEN. MICROBIOL. 139:1511-1522(1993).  
 RL [2]  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 25975;  
 RC MEDLINE; 92148377.  
 RA GIFFARD P.M., SIMPSON C.L., MILFARD C.P., JACQUES N.A.;  
 RT Molecular characterization of a cluster of at least two  
 RT glucosyltransferase genes in *streptococcus salivarius* ATCC 25975.;  
 RL J. GEN. MICROBIOL. 137:257-259(1991).  
 CC "Sequence of the gtf gene of *streptococcus salivarius*. and  
 evolution of the gtf genes of oral streptococci.";  
 RL J. GEN. MICROBIOL. 139:1511-1522(1993).  
 CC --!- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT  
 TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE  
 BECAUSE OF THEIR ABILITY TO ADHERER TO SMOOTH SURFACES AND MEDIATE  
 THE AGGRESSION OF BACTERIAL CELLS AND FOOD DEBRIS.  
 CC --!- CATALYTIC ACTIVITY: SUCROSE +  
 CC (1,6-ALPHA-D-GLUCOSYL)(N) = D-FRUCTOSE +  
 CC (1,6-ALPHA-D-GLUCOSYL)(N+).  
 CC --!- SUBCELLULAR LOCATION: EXTRACELLULAR.  
 CC --!- DISEASE: DENTAL CARIES.  
 CC --!- SIMILARITY: TO REGIONS OF BARLEY AND BACILLUS AMYLOLIQUEFACTENS  
 CC ALPHA AMYLASES AND RABBIT GLYCOGEN PHOSPHORYLASE.  
 DR EMBL; Z11872; G47531; -.  
 DR EMBL; 210873; G47528; -.  
 DR EMBL; M64111; G15350; -.  
 DR PFAM; PF00128; alpha-amylase; 1.  
 KW TRANSFERASE; GLYCOSYLTRANSFERASE; SIGNAL; REPEAT; DENTAL CARIES.  
 FT SIGNAL 1 42 POTENTIAL.  
 FT CHAIN 43 1599 GLYCOSYLTRANSFERASE S.  
 FT SIMILAR 867 967 TO ALPHA-AMYLASES.  
 FT SIMILAR 1132 1219 TO GLYCOGEN PHOSPHORYLASE.  
 SQ SEQUENCE 1599 AA; 176480 MW; 575C79C CRC32;  
 SP Query Match 68.4%; Score 106; DB 2; Length 1599;  
 ST Local Similarity 77.3%; Pred. No. 1 09e-09;  
 TC tches 17; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 Db 1402 TGAQWINGQHLYFDANGRQVKG 1423  
 Qy 1 TGARRINGOLLYFRANGQVKG 22

Search completed: Tue Jan 11 15:49:06 2000  
 Job time : 52 secs.



```
*****
*****
```



(TM)

```
*****
```

Release 3.1A John F. Collins, Biocomputing Research Unit.  
Copyright (c) 1993-1998 University of Edinburgh, U.K.

Distribution rights by Oxford Molecular Ltd

**MPSTCH\_PP** protein - protein database search, using Smith-Waterman algorithm

Run on: Tue Jan 11 15:37:52 2000; MasPar time 5.64 Seconds  
203.291 Million cell updates/sec

Title: >US-09-290-049-1

Description: (1-21) from US09290049.pep

Perfect Score: 148

Sequence: 1 ANDHISILEAWSNDTPYLHD 21

Scoring table: PAM 150  
Gap 15

Searched: 179066 seqs, 54579741 residues

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database: Trembl19

5:sp\_archaea 2:sp\_bacteria 3:sp\_fungi 4:sp\_human  
5:sp\_invertebrate 6:sp\_mammal 7:sp\_mhc 8:sp\_organelle  
9:sp\_phage 10:sp\_plant 11:sp\_rhodent 12:sp\_uniclassified  
13:sp\_vertebrate 14:sp\_virus

Statistics: Mean 28.660; Variance 41.580; scale 0.689

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

| Plt No. | Score | Query | Length | DB | ID     | Description              | Pred. No. |
|---------|-------|-------|--------|----|--------|--------------------------|-----------|
| 1       | 148   | 100.0 | 1390   | 2  | 069385 | GLUCOSYLTTRANSFERASE-SI  | 8.69e-19  |
| 2       | 148   | 100.0 | 1455   | 2  | 069390 | GLUCOSYLTTRANSFERASE-SI  | 8.69e-19  |
| 3       | 148   | 100.0 | 1455   | 2  | 069391 | GLUCOSYLTTRANSFERASE-I   | 8.69e-19  |
| 4       | 148   | 100.0 | 1476   | 2  | 069381 | GLUCOSYLTTRANSFERASE-I   | 8.69e-19  |
| 5       | 148   | 100.0 | 1476   | 2  | 069384 | GLUCOSYLTTRANSFERASE-I   | 8.69e-19  |
| 6       | 148   | 100.0 | 1476   | 2  | 069396 | GLUCOSYLTTRANSFERASE-I   | 8.69e-19  |
| 7       | 148   | 100.0 | 1476   | 2  | 069387 | GLUCOSYLTTRANSFERASE-I   | 8.69e-19  |
| 8       | 136   | 91.9  | 1455   | 2  | 069388 | GLUCOSYLTTRANSFERASE-SI  | 7.09e-16  |
| 9       | 136   | 91.9  | 1455   | 2  | 069382 | GLUCOSYLTTRANSFERASE-SI  | 7.09e-16  |
| 10      | 136   | 91.9  | 1455   | 2  | 055263 | GLUCOSYLTTRANSFERASE GT  | 1.11e-14  |
| 11      | 131   | 88.5  | 1590   | 2  | 059983 | GLUCOSYLTTRANSFERASE-I   | 1.11e-14  |
| 12      | 131   | 88.5  | 1590   | 2  | 059983 | GLUCOSYLTTRANSFERASE-S   | 8.76e-08  |
| 13      | 101   | 68.2  | 1462   | 2  | 069398 | GLUCOSYLTTRANSFERASE-S   | 8.76e-08  |
| 14      | 101   | 68.2  | 1462   | 2  | 069386 | GLUCOSYLTTRANSFERASE-S   | 8.76e-08  |
| 15      | 101   | 68.2  | 1462   | 2  | 069399 | GLUCOSYLTTRANSFERASE-S   | 8.76e-08  |
| 16      | 101   | 68.2  | 1462   | 2  | 069392 | GLUCOSYLTTRANSFERASE-S   | 8.76e-08  |
| 17      | 101   | 68.2  | 1462   | 2  | 069383 | GLUCOSYLTTRANSFERASE-S   | 8.76e-08  |
| 18      | 101   | 68.2  | 1508   | 2  | 052244 | GLUCOSYLTTRANSFERASE (E) | 1.78e-06  |
| 19      | 95    | 64.2  | 1577   | 2  | 054178 | GLUCOSYLTTRANSFERASE     | 1.78e-06  |
| 20      | 86    | 58.1  | 1449   | 2  | 068542 | GLUCOSYLTTRANSFERASE N   | 1.41e-04  |

| RESULT ID | 069385   | PRELIMINARY; | PRT; | 1390 AA. |
|-----------|--|--------------|------|----------|
| ID        | 069385;  |              |      |          |
| DT        | 01-AUG-1998 ('TREMBREL. 07, CREATED)   |              |      |          |
| DT        | 01-AUG-1998 ('TREMBREL. 07, LAST SEQUENCE UPDATE)  |              |      |          |
| DT        | 01-NOV-1998 ('TREMBREL. 08, LAST ANNOTATION UPDATE)  |              |      |          |
| DE        | GLYCOSYLTRANSFERASE-SI   |              |      |          |
| GN        | GTFC.  |              |      |          |
| RN        | [1]  |              |      |          |
| RP        | SEQUENCE FROM N.A.   |              |      |          |
| RC        | STRAIN=M74245;   |              |      |          |
| RY        | MEDLINE; 98231643.   |              |      |          |
| RA        | FUJIMURA T., TERRAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBOE S., KIMURA S., HAMADA S.;   |              |      |          |
| RT        | "Molecular analyses of glucosyltransferase genes among strains of <i>Streptococcus mutans</i> ;"   |              |      |          |
| RT        | FEMS MICROBIOL LETT. 161:331-336(1998).  |              |      |          |
| RL        | EMBL; D88655; D1027046; -.   |              |      |          |
| DR        | EMBL; D88655; D1027046; -.   |              |      |          |
| KW        | SEQUENCE, 1390 AA; 155375 MW; C211ETB8 CRC32;  |              |      |          |
| SQ        | Query Match: Best Local Similarity 100.0%; Score 148; DB: 2; length 1390; Matches: 21; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0; |              |      |          |
| DB        | 507 ANDHISILEAWSNDTPYLHD 527   |              |      |          |
| QY        | 1 ANDHISILEAWSNDTPYLHD 21  |              |      |          |
| RESULT ID | 069391   | PRELIMINARY; | PRT; | 1455 AA. |
| ID        | 069391;  |              |      |          |
| AC        | 069391;  |              |      |          |
| DT        | 01-AUG-1998 ('TREMBREL. 07, CREATED)   |              |      |          |
| DT        | 01-AUG-1998 ('TREMBREL. 07, LAST SEQUENCE UPDATE)  |              |      |          |
| DT        | 01-NOV-1998 ('TREMBREL. 08, LAST ANNOTATION UPDATE)  |              |      |          |
| DE        | GLUCOSYLTTRANSFERASE-SI.   |              |      |          |
| GN        | GTFC.  |              |      |          |
| RN        | [1]  |              |      |          |
| RP        | SEQUENCE FROM N.A.   |              |      |          |
| RC        | STRAIN=M74245;   |              |      |          |
| RY        | MEDLINE; 98231643.   |              |      |          |
| RA        | FUJIMURA T., TERRAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBOE S., KIMURA S., HAMADA S.;   |              |      |          |
| RT        | "Molecular analyses of glucosyltransferase genes among strains of <i>Streptococcus mutans</i> ;"   |              |      |          |
| RT        | FEMS MICROBIOL LETT. 161:331-336(1998).  |              |      |          |
| RL        | EMBL; D88655; D1027046; -.   |              |      |          |
| DR        | EMBL; D88655; D1027046; -.   |              |      |          |
| KW        | SEQUENCE, 1390 AA; 155375 MW; C211ETB8 CRC32;  |              |      |          |
| SQ        | Query Match: Best Local Similarity 100.0%; Score 148; DB: 2; length 1390; Matches: 21; Conservative: 0; Mismatches: 0; Indels: 0; Gaps: 0; |              |      |          |
| DB        | 507 ANDHISILEAWSNDTPYLHD 527   |              |      |          |
| QY        | 1 ANDHISILEAWSNDTPYLHD 21  |              |      |          |

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-MT251;  
 RX MEDLINE; 98231543.  
 RA FUJWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S.,  
 RA KIMURA S., HAMADA S.;  
 RT "Molecular analyses of glucosyltransferase genes among strains of  
 RT streptococci mutans.", FEMS MICROBIOL. LETT. 161:331-336(1998).  
 RL FEMS MICROBIOL. LETT. 161:331-336(1998).  
 DR EMBL; D88661; D1027054; -.  
 DR TRANSFERASE.  
 KW SEQUENCE; 1455 AA; 162804 MW; DB730514 CRC32;

Query Match 100.0%; Score 148; DB 2; Length 1455;  
 Best Local Similarity 100.0%; Pred. No. 8.69e-19;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 507 ANDHISILEAWSDNDDTPYLHD 527  
 1 ANDHISILEAWSDNDDTPYLHD 21

RESULT 3  
 ID 069390 PRELIMINARY; PRT; 1476 AA.  
 AC 069390;  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE GLUCOSYLTTRANSFERASE-I.  
 GN GTFB.  
 OS STREPTOCOCCUS MUTANS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-MT245;  
 RX MEDLINE; 98231643.  
 RA FUJWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S.,  
 RA KIMURA S., HAMADA S.;  
 RT "Molecular analyses of glucosyltransferase genes among strains of  
 RT streptococci mutans.", FEMS MICROBIOL. LETT. 161:331-336(1998).  
 RL FEMS MICROBIOL. LETT. 161:331-336(1998).  
 DR EMBL; D88660; D1027053; -.  
 KW TRANSFERASE.  
 SQ SEQUENCE 1476 AA; 165833 MW; 9119C72B CRC32;

Query Match 100.0%; Score 148; DB 2; Length 1476;  
 Best Local Similarity 100.0%; Pred. No. 8.69e-19;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 481 ANDHISILEAWSDNDDTPYLHD 501  
 Qy 1 ANDHISILEAWSDNDDTPYLHD 21

RESULT 4  
 ID 069381 PRELIMINARY; PRT; 1476 AA.  
 AC 069381;  
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE GLUCOSYLTTRANSFERASE-I.  
 GN GTFB.  
 OS STREPTOCOCCUS MUTANS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-MT8148;  
 RX MEDLINE; 98231643.  
 RA FUJWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S.,  
 RA KIMURA S., HAMADA S.;  
 RT "Molecular analyses of glucosyltransferase genes among strains of  
 RT streptococci mutans.", FEMS MICROBIOL. LETT. 161:331-336(1998).  
 RL FEMS MICROBIOL. LETT. 161:331-336(1998).  
 DR EMBL; D89977; D1027059; -.  
 KW TRANSFERASE.  
 SQ SEQUENCE 1476 AA; 165647 MW; 258D4ADD CRC32;

Query Match 100.0%; Score 148; DB 2; Length 1476;

RN RT STREPTOCOCCUS MUTANS.;  
 RP RL FEMS MICROBIOL. LETT. 161:331-336(1998).  
 RC DR EMBL; D88651; D1027041; -.  
 RX KW TRANSFERASE.

Query Match 100.0%; Score 148; DB 2; Length 1476;  
 Best Local Similarity 100.0%; Pred. No. 8.69e-19;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 481 ANDHISILEAWSDNDDTPYLHD 501  
 Qy 1 ANDHISILEAWSDNDDTPYLHD 21

RESULT 5  
 ID 069394 PRELIMINARY; PRT; 1476 AA.  
 AC 069384;  
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE GLUCOSYLTTRANSFERASE-I.  
 GN GTFB.  
 OS BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-MT245;  
 RX MEDLINE; 98231643.  
 RA FUJWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S.,  
 RA KIMURA S., HAMADA S.;  
 RT "Molecular analyses of glucosyltransferase genes among strains of  
 RT streptococci mutans.", FEMS MICROBIOL. LETT. 161:331-336(1998).  
 RL FEMS MICROBIOL. LETT. 161:331-336(1998).  
 DR EMBL; D88654; D1027045; -.  
 KW TRANSFERASE.  
 SQ SEQUENCE 1476 AA; 165792 MW; C76D8483 CRC32;

Query Match 100.0%; Score 148; DB 2; Length 1476;  
 Best Local Similarity 100.0%; Pred. No. 8.69e-19;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 481 ANDHISILEAWSDNDDTPYLHD 501  
 Qy 1 ANDHISILEAWSDNDDTPYLHD 21

RESULT 6  
 ID 069395 PRELIMINARY; PRT; 1476 AA.  
 AC 069395;  
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE GLUCOSYLTTRANSFERASE-I.  
 GN GTFB.  
 OS STREPTOCOCCUS MUTANS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-MT467;  
 RX MEDLINE; 98231643.  
 RA FUJWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S.,  
 RA KIMURA S., HAMADA S.;  
 RT "Molecular analyses of glucosyltransferase genes among strains of  
 RT streptococci mutans.", FEMS MICROBIOL. LETT. 161:331-336(1998).  
 RL FEMS MICROBIOL. LETT. 161:331-336(1998).  
 DR EMBL; D89977; D1027059; -.  
 KW TRANSFERASE.

Query Match 100.0%; Score 148; DB 2; Length 1476;



DE GLUCOSYLTRANSFERASE\_GTF-I.  
 OS  
 OC STREPTOCOCCUS\_SORBIUS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 RN STREPTOCOCCUS.  
 RP SEQUENCE FROM N.A.  
 RA STRAIN-MT347;  
 RC STRAIN-MT447;  
 RK MEDLINE; 9823163.  
 RA FUTIWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S.,  
 RA SATO S.;  
 RA KIMURA S., HAMADA S.;  
 RT "Molecular analyses of glucosyltransferase genes among strains of  
 Streptococcus mutans";  
 RL FEMS MICROBOL. LETT. 161:331-336(1998).  
 DR EMBL; D89919; D1027061; -.  
 DR PFAM; PF00128; alpha-amylase; 1.  
 KW TRANSFERASE.  
 SQ SEQUENCE 1590 AA; 176057 MW; 47186F6A CRC32;  
 Query Match 88.5%; Score 111; DB 2; Length 1590;  
 Best Local Similarity 85.7%; Pred. No. 1.1le-14;  
 Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 ANDHLSILEAWSNDNTPYLHD 21

RESULT 12 PRELIMINARY; PRT; 1590 AA.  
 ID 059983; PRT; 1590 AA.  
 AC 059983;  
 DT 01-NOV-1996 (TREMBLREL. 01, CREATED)  
 DT 01-NOV-1996 (TREMBLREL. 01, LAST SEQUENCE UPDATE)  
 DT 01-JAN-1999 (TREMBLREL. 09, LAST ANNOTATION UPDATE)  
 DE GLUCOSYLTRANSFERASE-I PRECURSOR (EC 2.4.1.5) (DEXTRANSUCRASE) (SUCROSE  
 6-GLUCOSYLTRANSFERASE).  
 DE GTF-I.  
 OS STREPTOCOCCUS\_SORBIUS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.  
 RN [1] SEQUENCE FROM N.A.  
 RC STRAIN-OMZ176;  
 RX MEDLINE; 94146405.  
 RA SATO S., INOUE M., HANADA N., AIZAWA Y., ISOBE Y., KATAYAMA T.;  
 RT "DNA sequence of the glucosyltransferase gene of serotype d".  
 RT Streptococcus sobrinus.";  
 RL DNA SEQ. 4:19-27(1993).  
 CC -1- CATALYTIC ACTIVITY: SUCROSE + (1,6-ALPHA-D-GLUCOSYL)(N+1) -  
 DR D-FRUCTOSE + (1,6-ALPHA-D-GLUCOSYL)(N+1).  
 EMBL; D13856; D1003482; -.  
 DR PFAM; PF00128; alpha\_amylase; 1.  
 DR SIGNAL; TRANSFERASE; GLYCOSYLTRANSFERASE.  
 F1 SIGNAL 1 38 POTENTIAL.  
 F1 CHAIN 39 1590 GLUCOSYLTRANSFERASE-I.  
 SQ SEQUENCE 1590 AA; 175955 MW; FE6A57D7 CRC32;

Query Match 88.5%; Score 131; DB 2; Length 1590;  
 Best Local Similarity 85.7%; Pred. No. 1.1le-14; Matches 18; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

Db 477 ANNHSVSIWAEWSNDNTPYLHD 497  
 QY 1 ANDHLSILEAWSNDNTPYLHD 21

RESULT 13 PRELIMINARY; PRT; 1462 AA.  
 ID 069398  
 AC 069398;  
 DT 01-AUG-1998 (TREMBLREL. 07, CREATED)  
 DT 01-AUG-1998 (TREMBLREL. 07, LAST SEQUENCE UPDATE)  
 DT 01-NOV-1998 (TREMBLREL. 08, LAST ANNOTATION UPDATE)  
 DE GLUCOSYLTRANSFERASE-S.  
 DE GTFD.  
 GN GTFD.  
 OS STREPTOCOCCUS MUTANS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.  
 RN [1] SEQUENCE FROM N.A.  
 RC STRAIN-MT4239;  
 RX MEDLINE; 9823163.  
 RA FUTIWARA T., TERAO Y., HOSHINO T., KAWABATA S., OOSHIMA T., SOBUE S.,  
 RA KIMURA S., HAMADA S.;  
 RT "Molecular analyses of glucosyltransferase genes among strains of  
 Streptococcus mutans";

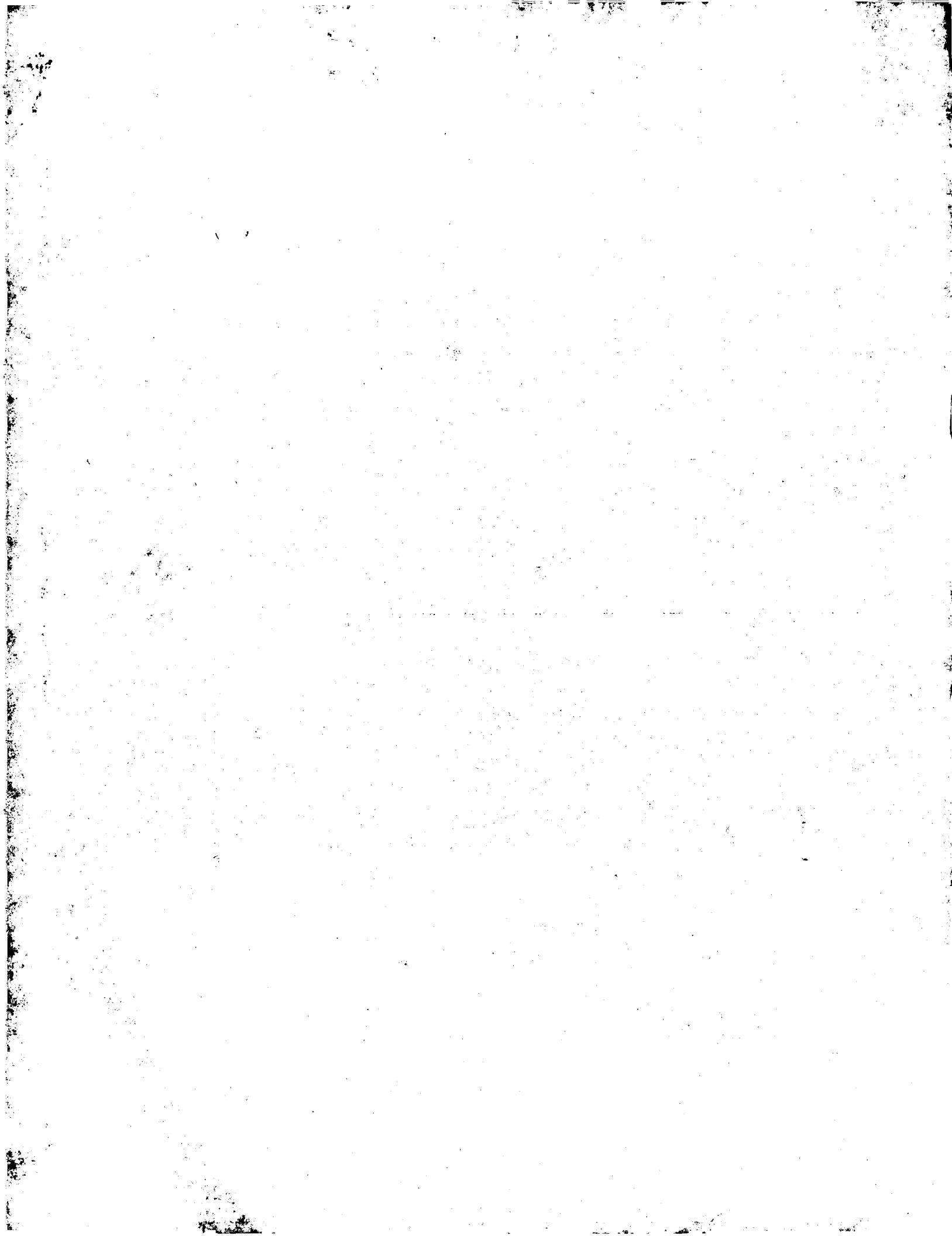
Thu Jan 13 09:13:35 2000

US-09-290-049-1.rspt

Page 5

RL FEMS MICROBIOL. LETT. 161:331-336(1998).  
DR EMBL; D88659; D1027051; -  
SQ SEQUENCE 1462 AA; 163554 MW; B057EF28 CRC32;  
Query Match 68 2%; Score 101; DB 2; Length 1462;  
Best Local Similarity 77.8%; Prcd. No. 8.76e-08;  
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
Db 495 AINHISILEAWSNDPQY 512  
Qy | :||| ||| ||| ||| |  
1 ANDHISILEAWSNDTPY 18

Search completed: Tue Jan 11 15:38:39 2000  
Job time : 47 secs.





Query Match 100.0%; Score 18; DB 1; Length 1475;  
 Best Local Similarity 100.0%; Pred. No. 1; 49e-18; 0; Mismatches 0; Indels 0; Gaps 0;  
 Matches 21; Conservative 21; Mismatches 0; Indels 0; Gaps 0;

Db 481 ANDHSILEAWSDNDTPYLHD 501  
 1 ANDHSILEAWSDNDTPYLHD 21

Query Match 100.0%; Score 18; DB 1; Length 1475;  
 Best Local Similarity 100.0%; Pred. No. 1; 49e-18; 0; Mismatches 0; Indels 0; Gaps 0;  
 Matches 21; Conservative 21; Mismatches 0; Indels 0; Gaps 0;

Db 507 ANDHSILEAWSYNTPYLHD 527  
 1 ANDHSILEAWSDNDTPYLHD 21

Query Match 91.9%; Score 136; DB 1; Length 1375;  
 Best Local Similarity 95.2%; Pred. No. 1; 1e-15; 0; Mismatches 1; Indels 0; Gaps 0;  
 Matches 20; Conservative 20; Mismatches 1; Indels 0; Gaps 0;

Db 507 ANDHSILEAWSYNTPYLHD 527  
 1 ANDHSILEAWSDNDTPYLHD 21

Query Match 91.9%; Score 136; DB 1; Length 1375;  
 Best Local Similarity 95.2%; Pred. No. 1; 1e-15; 0; Mismatches 1; Indels 0; Gaps 0;  
 Matches 20; Conservative 20; Mismatches 1; Indels 0; Gaps 0;

DR PIR: C33135; C33135.  
 DR PFAM: PF00128; alpha-amylase; 1.  
 DR TRANSFERASE; GLUCOSYLTRANSFERASE; SIGNAL; REPEAT; DENTAL CARIES.  
 KW  
 FT SIGNAL\_ 1 34  
 FT CHAIN 35 1375 . GLUCOSYLTRANSFERASE-SI.  
 FT DOMAIN 35 1050 . CATALYTIC (APPROXIMATE).  
 FT DOMAIN 1126 1375 GLUCAN-BINDING (APPROXIMATE).  
 FT DOMAIN 1125 1375 2.4 A, 1 C AND 1 AC REPEATS.  
 FT REPEAT 1126 1159 A REPEAT.  
 FT REPEAT 1169 1200 A REPEAT.  
 FT REPEAT 1227 1238 C REPEAT.  
 FT REPEAT 1253 1303 AC REPEAT.  
 FT REPEAT 1318 1330 A REPEAT (INCOMPLETE).  
 PRT; 1375 AA.

RESULT 2  
 ID GTFC STRMU STANDARD;  
 AC P13470; P05427;  
 DT 01-NOV-1988 (REL. 09, CREATED)  
 DT 01-JAN-1990 (REL. 13, LAST ANNOTATION UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE GLUCOSYLTRANSFERASE-SI PRECURSOR (EC 2.4.1.5) (GTF-SI)  
 DE (DEXTRANSUCRASE).  
 GN GTFC.  
 OS STREPTOCOCCUS MUTANS.  
 OC BACTERIA; FIRMITICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAINS-G5;  
 RX MEDLINE; 89137980.  
 RA UEDA S., SHIROZA T., KURAMITSU H.K.;  
 RT "Sequence analysis of the gtfc gene from Streptococcus mutans GS-5.";  
 RL GENB 69:101-109(1988).  
 RN [2]  
 RP SEQUENCE OF 1-349 FROM N.A.  
 RC STRAINS-G5;  
 RX MEDLINE; 8730813.  
 RA SHIROZA T., UEDA S., KURAMITSU H.K.;  
 RT "Sequence analysis of the gtfc gene from Streptococcus mutans.";  
 RL J. BACTERIOL. 169:4263-4270(1987).  
 CC - FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE AGGRESSION OF BACTERIAL CELLS AND FOOD DEBRIS.  
 CC - CATALYTIC ACTIVITY: SUCROSE + (1,6-ALPHA-D-GLUCOSYL) (N) - D-FRUCTOSE + (1,6-ALPHA-D-GLUCOSYL) (N+1).  
 CC - SUBCELLULAR LOCATION: SECRETED.  
 CC - DISEASE: DENTAL CARIES.  
 CC - DISEASE: DENTAL CARIES.  
 CC - GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH FORMS OF GLUCANS.  
 CC - SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: M22054; GL153643; -.  
 CC DR M22054; GL153643; -.  
 DR EMBL: M17361; GL153641; -.  
 DR PIR: JY0345; JT0345.

CC DR PIR: C33135; C33135.  
 DR PFAM: PF00128; alpha-amylase; 1.  
 DR TRANSFERASE; GLUCOSYLTRANSFERASE; SIGNAL; REPEAT; DENTAL CARIES.  
 KW  
 FT SIGNAL\_ 1 34  
 FT CHAIN 35 1375 . GLUCOSYLTRANSFERASE-SI.  
 FT DOMAIN 35 1050 . CATALYTIC (APPROXIMATE).  
 FT DOMAIN 1126 1375 GLUCAN-BINDING (APPROXIMATE).  
 FT DOMAIN 1125 1375 2.4 A, 1 C AND 1 AC REPEATS.  
 FT REPEAT 1126 1159 A REPEAT.  
 FT REPEAT 1169 1200 A REPEAT.  
 FT REPEAT 1227 1238 C REPEAT.  
 FT REPEAT 1253 1303 AC REPEAT.  
 PRT; 1592 AA.

RESULT 3  
 ID GTF2\_STRDO STANDARD;  
 AC P27471;  
 DT 01-AUG-1992 (REL. 23, CREATED)  
 DT 01-AUG-1992 (REL. 23, LAST ANNOTATION UPDATE)  
 DT 15-DEC-1998 (REL. 37, LAST ANNOTATION UPDATE)  
 DE GLUCOSYLTRANSFERASE-I PRECURSOR (EC 2.4.1.5) (GTF-I) (DEXTRANSUCRASE)  
 DE (SUCROSE 6-GLUCOSYLTRANSFERASE).  
 OS STREPTOCOCCUS DOMINI (STREPTOCOCCUS SOBRINUS).  
 OC BACTERIA; FIRMITICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC STREPTOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=6715;  
 RX MEDLINE; 91123227.  
 RA ABO H., MATSUMURA T., KODAMA T., OHTA H., FUKUI K., KATO K.,  
 RA KAGAWA H.;  
 RT "Peptide sequences for sucrose splitting and glucan binding within Streptococcus sobrinus glucosyltransferase (water-insoluble glucan synthetase)."; RT  
 RL J. BACTERIOL. 173:989-996(1991).  
 CC - FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE AGGRESSION OF BACTERIAL CELLS AND FOOD DEBRIS.  
 CC - CATALYTIC ACTIVITY: SUCROSE + (1,6-ALPHA-D-GLUCOSYL) (N) - D-FRUCTOSE + (1,6-ALPHA-D-GLUCOSYL) (N+1).  
 CC - SUBCELLULAR LOCATION: SECRETED.  
 CC - DISEASE: DENTAL CARIES.  
 CC - GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED GLUCOSE AND SOME 1,6 LINKAGES), GTF-S SYNTHESIZES WATER-SOLUBLE GLUCANS (ALPHA 1,6-GLUCOSE). GTF-SI SYNTHESIZES BOTH FORMS OF GLUCANS.  
 CC - SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-BINDING PROTEIN FROM S. MUTANS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC EMBL: D90213; G317033;  
 CC DR PIR: A38175; A38175.  
 DR PFAM: PF00128; alpha-amylase; 1.  
 DR HSSP: P00695; 2HEE.  
 DR TRANSFERASE; GLUCOSYLTRANSFERASE; SIGNAL; REPEAT; DENTAL CARIES.









This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

---

DR EMBL; X12948; G65028; -.

DR EMBL; S74063; G24126; -.

DR PIR; S01930; TXVLRF.

DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.

DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.

DR PROSITE; PS00479; DAG\_PE\_BINDING\_DOMAIN; 1.

DR PROSITE; PS00011; PROTEIN\_KINASE\_DOM; 1.

DR PFAM; PF00059; kinase; 1.

DR PFAM; PF00130; DAG\_PE-bind; 1.

DR HSSP; P04049; 1FAC.

KW TRANSFERASE; SERINE/THREONINE-PROTEIN KINASE; PROTO-ONCOGENE; ZINC; ATP-BINDING; PHORBOL-ESTER BINDING.

FT DOMAIN 138 183 PHORBOL-ESTER AND DAG BINDING.

FT DOMAIN 340 600 PROTEIN KINASE.

FT NP\_BIND 345 354 ATP (BY SIMILARITY).

FT BINDING 366 366 ATP (BY SIMILARITY).

FT ACT\_SITE 459 459 BY SIMILARITY.

FT CONFFLICT 309 309 K->R (IN REF. 2).

SQ 638 AA; 71959 MW; A46048CC CRC32;

Query Match 43.2%; Score 64; DB 1; Length 638; Best Local Similarity 30.0%; Pred. No. 1.94e+00; Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

---

RESULT 13 ID PMA1\_DUNBI STANDARD; PRT; 1131 AA.

DB 404 TKDNLAIVTQWCGSSLVYH 423 AC P24211; DT 01-OCT-1996 (REL. 34, CREATED)

DT 01-OCT-1995 (REL. 34, LAST SEQUENCE UPDATE)

DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE PLASMA MEMBRANE ATPase (EC 3.6.1.35) (PROTON PUMP).

GN PMA1.

OS DUNALIELLA BIOCULATA.

OC EUKARYOTA; VIRIDPLANTAE; CHLOROPHYTA; CHLOROPHYCEAE; VOLVOCALES; DINALIELLACEAE; DUNALIELLA.

[1] SEQUENCE FROM N.A.

RX STRAIN-SRG 19-4;

RX MEDLINE: 95315230.

RA WOLF A.H., SLAYMAN C.W., GRADMANN D.; RT Primary structure of the plasma membrane H(+) -ATPase from the halotolerant alga Dunaliella bioculata.;

RL PLANT BIOL. 28:557-566(1995).

CC "1- FUNCTION: THE PLASMA MEMBRANE ATPASE OF PLANTS AND FUNGI IS A HYDROGEN ION PUMP. THE PROTON GRADIENT IT GENERATES DRIVES THE ACTIVE TRANSPORT OF NUTRIENTS BY H+-SYMPORT. THE RESULTING EXTERNAL ACIDIFICATION AND/OR INTERNAL ALKINIZATION MAY MEDIATE GROWTH RESPONSES.

CC "- CATALYTIC ACTIVITY: ATP + H(2)O = ADP + ORTHOPHOSPHATE.

CC "- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN.

CC "- SIMILARITY: BELONGS TO THE CATION TRANSPORT ATPASES FAMILY (EL-E2 ATPases).

---

RESULT 14 ID KRAF\_MSV76 STANDARD; PRT; 323 AA.

DB 774 NDQJISVLQWCEDE 787 AC P00532; DT 21-JUL-1986 (REL. 01, CREATED)

DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)

DT 01-NOV-1995 (REL. 32, LAST ANNOTATION UPDATE)

DE RAF SERINE/THREONINE-PROTEIN KINASE TRANSFORMING PROTEIN (BC 2.7.1.-).

GN V-RAF.

OS MURINE SARCOMA VIRUS 3611.

OC VIRUSES; RETROID VIRUSES; RETROVIRIDAE; MAMMALIAN TYPE C RETROVIRUSES.

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE: 84121298.

RA KAN N.C., FLORENDLIS C.S., MARK G.E., DUESBERG P.H., PAPAS T.S.; RT "A common onc gene sequence transduced by avian carcinoma virus MH2 and by murine sarcoma virus 3611.";

RT SCIENCE 223:813-816(1984).

RL SCIENCE 223:813-816(1984). [2]

RP SEQUENCE FROM N.A.

RX MEDLINE: 84177180.

RA MARK G.E., RAPP U.R.; RT MARK G.E., RAPP U.R.;

RT "Primary structure of v-raf: relatedness to the src family of oncogenes";

RL SCIENCE 224:295-298(1984).

CC "- THIS PROTEIN IS SYNTHESIZED AS A GAG-RAF POLYPROTEIN.

CC "- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-PROTEIN KINASES. BELONGS TO THE RAF SUBFAMILY.

---

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

---

DR EMBL; K01691; G332207; ALT\_INMT.

DR PIR; A00538; TWMVF6.

DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.

DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.

DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.

DR PFAM; PF00069; kinase; 1.

DR HSSP; P11362; 1FG1.

KW POLYPROTEIN; TRANSFORMING PROTEIN; SERINE/THREONINE-PROTEIN KINASE;  
 KW TRANSFERASE; ONCOGENE; ATP-BINDING.  
 FT DOMAIN 24 284 PROTEIN KINASE.  
 FT NP\_BIND 30 38 ATP (BY SIMILARITY).  
 FT BINDING 50 50 ATP (BY SIMILARITY).  
 FT ACT\_SITE 143 143 BY SIMILARITY.  
 SQ SEQUENCE 323 AA: 36883 MW: B3FF54BE CRC32:  
 Query Match 41.9%; Score 62; DB 1; Length 380;  
 Best Local Similarity 30.0%; Pred. No. 4.23e+00;  
 Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;  
 Db 88 TKDNLATVQWSEGGSSLYKH 107  
 Qy 1 ANDHLSILEAWSNDTPYLH 20

RESULT 15  
 KMIL\_AVIMH STANDARD: PRM: 380 AA.  
 P00531;  
 D1 21-JUL-1986 (REL. 01, CREATED)  
 DT 21-JUL-1986 (REL. 01, LAST SEQUENCE UPDATE)  
 DE 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  
 DE MIL SERINE/THREONINE PROTEIN KINASE TRANSFORMING PROTEIN  
 DE (BC 2.7.1.-).  
 GN V-MIL OR V-HMT.  
 OS AVIAN RETROVIRUS MH2.  
 OC VIRUSES; RETROID VIRUSES; RETROVIRIDAE; AVIAN TYPE C RETROVIRUSES.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 8419151.  
 RA SUTRAVE P., BONNER T.I., RAPP U.R., JANSEN H.W., PATSCHINSKY T.,  
 RA BISTER K.;  
 RT "Nucleotide sequence of avian retroviral oncogene v-mil: homologue of  
 murine retroviral oncogene v-raf.";  
 RL NATURE 309:85-88(1984).  
 12.]  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE; 84121298.  
 RA KAN N.C., FLORDELLIS C.S., MARK G.E., DUESBERG P.H., PAPAS T.S.;  
 RT "A common onc gene sequence transduced by avian carcinoma virus MH2  
 RL and by murine sarcoma virus 3611.";  
 CC SCIENCE 223:813-816(1984).  
 -I- DISEASE: BY ITSELF THE V-MIL ONCOGENE HAS ONLY WEAK TRANSFORMING  
 CAPACITY BUT IT ABOLISHES THE GROWTH FACTOR REQUIREMENTS OF AVIAN  
 MACROPHAGES TRANSFORMED BY OTHER ONCOGENES.  
 CC -I- THIS PROTEIN IS SYNTHESIZED AS A GAG-MIL OR GAG-MIL POLYPROTEIN.  
 -I- SIMILARITY: WITH THE CONSERVED CATALYTIC DOMAINS OF SER/THR-  
 PROTEIN KINASES. BELONGS TO THE MIL/RAF SUBFAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC  
 EMBL; X00534; E235847; ALT\_INIT.  
 DR PIR; A00639; TVFNM.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PFAM; PF00059; Proteinase; 1.  
 DR HSSP; P11362; IFG1.  
 DR EMBL; X00534; E235847; ALT\_INIT.  
 DR PROSITE; PS00107; PROTEIN\_KINASE\_ATP; 1.  
 DR PROSITE; PS00108; PROTEIN\_KINASE\_ST; 1.  
 DR PFAM; PF00059; Proteinase; 1.  
 DR HSSP; P11362; IFG1.  
 KW POLYPROTEIN; TRANSFORMING PROTEIN; SERINE/THREONINE-PROTEIN KINASE;  
 KW TRANSFERASE; ONCOGENE; ATP-BINDING.  
 FT DOMAIN 82 341 PROTEIN KINASE.  
 FT NP\_BIND 88 95 ATP (BY SIMILARITY).  
 FT BINDING 108 108 ATP (BY SIMILARITY).  
 FT ACT\_SITE 201 201 BY SIMILARITY.  
 FT CONFLICT 211 211 G -> E (IN REF. 2).  
 SQ SEQUENCE 380 AA: 42853 MW: 1D7B2AB4 CRC32:

Query Match 41.9%; Score 62; DB 1; Length 380;  
 Best Local Similarity 30.0%; Pred. No. 4.23e+00;  
 Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;  
 Db 146 TKDNLATVQWCEGGSSLYKH 165  
 Qy 1 ANDHLSILEAWSNDTPYLH 20

Search completed: Tue Jan 11 15:37:35 2000  
 Job time : 9 secs.



DATE 28-Aug-1992 #sequence\_revision 28-Aug-1992 #text\_change

ACCESSIONS A38175  
REFERENCE Abo, H.; Msumura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, K.; Kagaya, H. (1991) 173:989-996  
#authors J. Bacteriol. Peptide sequences for sucrose splitting and glucan binding within Streptococcus sobrinus glucosyltransferase.

#accession A38175  
#status preliminary  
#molecule\_type DNA  
#residues 1-1592 ##label ABO  
#cross-references GB:ID920213; NID:9217032; PID:d1014946; PID:9217033

FEATURE #superfamily cpl repeat homology  
#domain cpl repeat homology #label CPL1  
#domain cpl repeat homology #label CPL2  
#domain cpl repeat homology #label CPL3  
#domain cpl repeat homology #label CPL4  
#domain cpl repeat homology #label CPL5  
#domain cpl repeat homology #label CPL6  
#domain cpl repeat homology #label CPL7  
#length 1592 #molecular\_weight 176167 #checksum 5940

SUMMARY Query Match 72.3% Score 112; DB 2; Length 1592; Best Local Similarity 77.3%; Pred. No. 5.30e-10; Matches 17; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Db 1298 TGATGTCGKLYFKANGQVKG 1319  
Oy 1 TGARTINGQQLYFRANGVQVK 22

RESULT 3

ENTRY A44811 #type complete  
TITLE glucoyltransferase (EC 2.4.1.-) I - Streptococcus salivarius  
ORGANISM #formal\_name Streptococcus salivarius  
DATE 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change

ACCESSIONS A44811; S22726; S28809  
REFERENCE Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.  
#authors J. Gen. Microbiol. (1991) 137:2577-2593  
#Journal Molecular characterization of a cluster of at least two glucosyltransferase genes in Streptococcus salivarius ATCC 25975.

#cross-references EMBL:ZI1873; NID:947526; PID:947527  
#accession A44811  
#molecule\_type DNA  
#residues 1-1518 ##label GTF  
#cross-references MUID:92148377  
#note sequence extracted from NCBI backbone (NCBIN:81050, NCBIP:81052)

GENETICS #gtfJ  
CLASSIFICATION #superfamily cpl repeat homology  
KEYWORDS glycosyltransferase; hexosyltransferase

FEATURE #domain cpl repeat homology #label CPL4  
#length 1518 #molecular\_weight 167730 #checksum 1334

Query Match 70.3% Score 109; DB 2; Length 1518; Best Local Similarity 72.7%; Pred. No. 2.45e-09; Matches 16; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

Db 1383 TGAQVINGQHLYFNADGSQVKG 1404  
Oy 1 TGARTINGQQLYFRANGVQVK 22

RESULT 4

ENTRY JC5473 #type complete  
TITLE dextranucrase (EC 2.4.1.5) - Leuconostoc mesenteroides  
ORGANISM #formal\_name Leuconostoc mesenteroides  
DATE 07-Jul-1997 #sequence\_revision 29-Aug-1997 #text\_change

29-Aug-1997

ACCESSIONS JC5473  
REFERENCE Monchi, V.; Willemot, R.M.; Remaud-Simeon, M.; Croux, C.; Gene (1996) 182:23-32  
#authors Cloning and sequencing of a gene coding for a novel dextranucrase from Leuconostoc mesenteroides NRRL B-1299 synthetizing only alpha (1-6) and alpha (1-3) linkages.

#cross-references MUID:97136686

FEATURE #status nucleic acid sequence not shown; translation not shown  
#cross-references GB:ID38181

SUMMARY Query Match 78.87% Score 1290 ##label MON  
Best Local Similarity 63.6%; Pred. No. 6.75e-09; Matches 14; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Db 1181 TGDRIIKGONLYNQNGIQMKG 1202  
Oy 1 TGARTINGQQLYFRANGVQVK 22

RESULT 5

ENTRY JT0345 #type complete  
TITLE dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans (strain GS-5)

ALTERNATE\_NAMES sucrose 6-glucosyltransferase  
ORGANISM #formal\_name Streptococcus mutans  
DATE 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change

22-Nov-1996

ACCESSIONS JT0345; C33135  
REFERENCE Ueda, S.; Shiroza, T.; Kuramitsu, H.K.  
#authors Gene (1988) 69:101-109  
#journal Sequence analysis of the gtfC gene from Streptococcus mutans GS-5.

#cross-references MUID:88137930  
#accession JT0345  
#molecule\_type DNA  
#residues 1-1375 ##label UED  
#experimental\_source GS-5

REFERENCE A33135  
#authors Shiroza, T.; Ueda, S.; Kuramitsu, H.K.  
#journal J. Bacteriol. (1987) 169:4263-4270  
#title Sequence analysis of the gtfB gene from Streptococcus mutans MUID:87308013

#cross-references accession C33135  
#status preliminary  
#molecule\_type DNA  
#residues 1-349 ##label SHI  
#cross-references GB:ID17361

GENETICS #gene  
FUNCTION #gtfc  
#description catalyzes the synthesis of both water-soluble and

CLASSIFICATION #superfamily cpl repeat homology; glycosyltransferase; hexosyltransferase; water-insoluble glucans from glucose.

KEYWORDS duplication; glycosyltransferase; hexosyltransferase; transferase

FEATURE 1-34

35-1375

ENTRY 6

ORGANISM #domain signal sequence #status predicted #label SIG\#product glucohydrolase #status predicted #label

DATE MAY

ACCESSIONS #domain cpl repeat homology #label CPI\#domain cpl repeat homology #label CP2\#domain cpl repeat homology #label CP3

REFERENCE #length 1375 #molecular-weight 153021 #checksum 7015

SUMMARY Query Match 68.4%; Score 106; DB 2; Length 1375; Best Local Similarity 68.2%; Pred. No. 1.12e-08; Matches 15; Conservative 4; Mismatches 3; Indels 0; Gaps 0;

Db 1264 TGTIVVNGQRLYFKPKNVQAKG 1285

Qy 1 TGARTINGQOLLYFRANGVQVK 22

ENTRY 7

ORGANISM S22737 #type complete glucohydrolase (EC 2.4.1.-) S - Streptococcus salivarius

DATE 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change 09-Sep-1997

ACCESSIONS S22737; S28810; B44811; S22727

REFERENCE #authors Jacques, N. submitted to the EMBL Data Library, March 1992

#accession S22737

#molecule-type DNA #residues 1-1599 #label JAC

#cross-references EMBL:211877; NID:947530; PID:947531

REFERENCE A44811

#authors Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.

#journal J. Gen. Microbiol. (1991) 137:2577-2593

title Molecular characterization of a cluster of at least two glucosyltransferase genes in Streptococcus salivarius ATCC 25975

#cross-references MUID:92148377

#accession S28810

#molecule-type DNA #residues 1-151 #label GIF

#cross-references EMBL:211873

GENETICS #gene gtfK

CLASSIFICATION #superfamily cpl repeat homology

KEYWORDS glycosyltransferase; hexosyltransferase

FEATURE 181-201

1127-1146 #domain cpl repeat homology #label CPI\#domain cpl repeat homology #label CP3\#domain cpl repeat homology #label CP4\#domain cpl repeat homology #label CP8\#domain cpl repeat homology #label CP6\#domain cpl repeat homology #label CP7

1122-1211 1257-1276

1277-1297 1321-1340

1341-1351 1385-1404

SUMMARY #length 1431 #molecular-weight 159844 #checksum 4333

Query Match 68.4%; Score 106; DB 2; Length 1599; Best Local Similarity 77.3%; Pred. No. 1.12e-08; Matches 17; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db 1402 TGAQVQNGQHLYFDANGRQVK 1423

Qy 1 TGARTINGQOLLYFRANGVQVK 22

RESULT 7

ENTRY A41483 #type complete glucosyltransferase (EC 2.4.1.-) gtfS precursor -

ORGANISM #formal\_name Streptococcus sobrinus

DATE 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Sep-1997

ACCESSIONS A41483

REFERENCE #authors Gilmore, K.S.; Russell, R.R.B.; Ferretti, J.J.

#cross-references MUID:90316665

CLASSIFICATION #molecule-type DNA

KEYWORDS ##cross-references GB:M30943; NID:g153652; PID:g153653

REFERENCE A41483

Query Match 67.1%; Score 104; DB 2; Length 1365; Best Local Similarity 63.6%; Pred. No. 3.05e-08; Matches 14; Conservative 5; Mismatches 3; Indels 0; Gaps 0;

Db 1235 TGEQTIDGQKFFQDNQVQVK 1256

Qy 1 TGARTINGQOLLYFRANGVQVK 22

RESULT 8

ENTRY A4566 #type complete dextranucrase (EC 2.4.1.5) precursor - Streptococcus mutans

ORGANISM #formal\_name Streptococcus mutans

DATE 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 07-Jul-1995

ACCESSIONS A4566

REFERENCE A4566

#authors Honda, O.; Kato, C.; Kuramitsu, H. K.

#journal J. Gen. Microbiol. (1990) 136:2099-2105

#title Nucleotide sequence of the Streptococcus mutans gtfd gene encoding the glucosyltransferase-S enzyme.

CLASSIFICATION #superfamily cpl repeat homology

KEYWORDS glycosyltransferase; hexosyltransferase

FEATURE 181-201

1127-1146 #domain cpl repeat homology #label CPI\#domain cpl repeat homology #label CP3\#domain cpl repeat homology #label CP4\#domain cpl repeat homology #label CP8\#domain cpl repeat homology #label CP6\#domain cpl repeat homology #label CP7

1122-1211 1257-1276

1277-1297 1321-1340

1341-1351 1385-1404

SUMMARY #length 1431 #molecular-weight 159844 #checksum 4333

Query Match 60.0%; Score 93; DB 2; Length 1431; Best Local Similarity 59.1%; Pred. No. 6.72e-06; Matches 13; Conservative 4; Mismatches 5; Indels 0; Gaps 0;

Db 1332 TGSQTAGKKIFASDGQVK 1353

Qy 1 TGARTINGQOLLYFRANGVQVK 22

RESULT 9

ENTRY A37184 #type complete glucan-binding protein - Streptococcus mutans

ORGANISM #formal\_name Streptococcus mutans

DATE 31-Jan-1992 #sequence\_revision 31-Jan-1992 #text\_change 09-Sep-1997

ACCESSIONS A37184

REFERENCE #authors Bans, J.A.; Russell, R.R.B.; Ferretti, J.J.

#cross-references Infec. Immun. (1990) 58:667-673

#title sequence analysis of the gene for the glucan-binding protein  
of *Streptococcus mutans* Ingbrtt.  
#accession M0945; NID:9153637; PID:9153638  
#classification A3714  
#status preliminary  
#molecule\_type DNA  
##residues 1-563 ##label BAN  
CLASSIFICATION #superfamily cpl repeat homology  
FEATURE #cross-references GB:M3045; NID:9153637; PID:9153638  
#status preliminary  
#molecule\_type DNA  
##residues 1-563 ##label BAN  
#accession S55166  
REFERENCE S5612  
#authors Katsoulou, C.; Tzermia, M.; Alexandraki, D.  
#molecule\_type DNA  
##residues 1-382 ##label KAT  
#accession S56166  
REFERENCE S7143  
#authors Katsoulou, C.; Tzermia, M.; Tavernarakis, N.; Alexandraki, D.  
#molecule\_type DNA  
##residues 1-382 ##label KAW  
#accession S5629  
REFERENCE S7143  
#authors Yeast (1996) 12:787-797  
#molecule\_type DNA  
##residues 1-382 ##label KAF  
#accession S7150  
#status nucleic acid sequence not shown; translation not shown  
#molecule\_type DNA  
##residues 1-382 ##label KAF  
#note the nucleotide sequence was submitted to the EMBL Data  
reading frames including homologues of genes clustered on  
the right arm of chromosome XI.  
#cross-references EMBL:18731; NID:9854542; PID:9854550  
#note the nucleotide sequence was submitted to the EMBL Data  
Library, May 1995  
GENETICS  
#map\_position 10L  
#length 382 #molecular-weight 44862 #checksum 8833  
SUMMARY  
ORGANISM Query Match 51.0%; Score 79; DB 2; Length 382;  
DATE Best Local Similarity 38.9%; Pred. No. 4.46e-03;  
ACCESSIONS Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;  
REFERENCE A55221  
#authors Sun, J.W.; Wanda, S.Y.; Camilli, A.; Curtiss III, R.  
#title J. Bacteriol. (1994) 176:7213-7222  
Cloning and DNA sequencing of the dextranase inhibitor gene  
(dei), from *Streptococcus sobrinus*.  
#cross-references M0945; NID:95050304  
#accession A55221  
RESULT 10  
ENTRY A55221 #type complete  
TITLE destranase inhibitor precursor - *Streptococcus sobrinus*  
ORGANISM (strain UAB66, serotype g)  
DATE #formal\_name *Streptococcus sobrinus*  
23-Mar-1995 #sequence\_revision 23-Mar-1995 #text\_change  
17-Mar-1999  
ACCESSIONS A55221  
REFERENCE A55221  
#authors Sun, J.W.; Wanda, S.Y.; Camilli, A.; Curtiss III, R.  
#title J. Bacteriol. (1994) 176:7213-7222  
Cloning and DNA sequencing of the dextranase inhibitor gene  
(dei), from *Streptococcus sobrinus*.  
#cross-references M0945; NID:95050304  
#accession A55221  
RESULT 11  
ENTRY A55166 #type complete  
TITLE hypothetical protein YJL147c - yeast (*Saccharomyces cerevisiae*)  
ORGANISM hypothetical protein J0639  
DATE #formal\_name *Saccharomyces cerevisiae*  
08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change  
05-Jun-1998  
ACCESSIONS S55166; S56929; S71650  
REFERENCE S55159  
#authors Katsoulou, C.; Tzermia, M.; Alexandraki, D.  
#submission submitted to the EMBL Data Library, May 1995  
#description The complete sequence of a 40.7 kb segment located on the  
left arm of yeast chromosome X identified 13 known genes  
and revealed 13 new open reading frames including  
homologues to other yeast hypothetical proteins.  
#accession S55166  
#molecule\_type DNA  
##residues 1-382 ##label KAW  
#accession S56166  
REFERENCE S7143  
#authors Katsoulou, C.; Tzermia, M.; Tavernarakis, N.; Alexandraki, D.  
#molecule\_type DNA  
##residues 1-382 ##label KAF  
#accession S7150  
#status nucleic acid sequence not shown; translation not shown  
#molecule\_type DNA  
##residues 1-382 ##label KAF  
#note the nucleotide sequence was submitted to the EMBL Data  
reading frames including homologues of genes clustered on  
the right arm of chromosome XI.  
#cross-references EMBL:18731; NID:9854542; PID:9854550  
#note the nucleotide sequence was submitted to the EMBL Data  
Library, May 1995  
GENETICS  
#map\_position 10L  
#length 382 #molecular-weight 44862 #checksum 8833  
SUMMARY  
ORGANISM Query Match 51.0%; Score 79; DB 2; Length 382;  
DATE Best Local Similarity 38.9%; Pred. No. 4.46e-03;  
ACCESSIONS Matches 7; Conservative 7; Mismatches 4; Indels 0; Gaps 0;  
REFERENCE F71014  
#authors Hino, Y.; Yamamoto, S.; Sekine, M.; Baba, S.; Kosuri, H.;  
Kosoyama, A.; Nagai, Y.; Sakai, M.; Ogura, K.; Otsuka, R.;  
Nakazawa, H.; Takamiya, M.; Ohfuku, Y.; Funahashi, T.;  
Tanaka, T.; Kudo, Y.; Yamazaki, J.; Kushida, N.; Oguchi, A.;  
Aoki, K.; Yoshizawa, T.; Nakamura, Y.; Robb, F.T.;  
Horikoshi, K.; Masuchi, Y.; Shizuya, H.; Kikuchi, H.  
#journal DNA Res. (1998) 5:55-76  
#title Complete sequence and gene organization of the genome of a  
hyper-thermophilic archaeabacterium, *Pyrococcus horikoshii*  
#cross-references M0945; NID:98344137  
#accession F71014  
#status preliminary; nucleic acid sequence not shown;  
#molecule\_type DNA  
##residues 1-329 ##label KAW  
#cross-references GB:Ap00006; NID:93236133; PID:d1031461; PID:93257835  
#note this accession replaces an interim accession for a  
#note sequence replaced by GenBank  
GENETICS  
#note PH112  
FUNCTION

#description probably responsible for energy-coupling to the transport system  
 CLASSIFICATION superfamily inner membrane protein malk; ATP-binding cassette homology  
 KEYWORDS • ATP; binding; protein-dependent transport system; nucleotide binding; oligopeptide transport; p-loop

FEATURE 241 #domain ATP-binding cassette homology #label ABC\\  
 45-53 #region nucleotide-binding motif A (P-loop)\\  
 188-189 #region nucleotide-binding motif B

SUMMARY #length 329 #molecular-weight 36504 #checksum 6349

Query Match Best Local Similarity 44.4%; Pred. No. 1.08e-02; Length 329;  
 Matches 8; Conservative 5; Mismatches 5; Indels 0; Gaps 0;

Db 66 GRIVGQVLVYREDGSVY 83  
 Qy 3 ARTINGQOLLYFRANGVQV 20

ENTRY 13  
 TITLE S40016 #type complete  
 ORGANISM doc protein - phage PI  
 DATE #formal\_name Phage PI  
 25-Dec-1994 #sequence\_revision 17-May-1996 #text\_change  
 20-Jan-1999

ACCESSIONS S40016; S38555

REFERENCE #authors Lehnherr, H.; Maguin, E.; Jafri, S.; Yarmolinsky, M.B.  
 #journal J. Mol. Biol. (1993) 233:414-428  
 Plasmid addiction genes of bacteriophage PI: doc, which causes cell death on curing of prophage, and pid, which prevents host death when prophage is retained.

#cross-references MVID:94016561  
 #accession S40016  
 #molecule-type DNA  
 #residues 1-126 #label LEH  
 #cross-references GB:M95666; NID:9463276; PID:9463277

REFERENCE S3853  
 #authors Schmidt, C.; Lehnherr, H.; Guidolin, A.; Abre, W.  
 #submission submitted to the EMBL Data Library, November 1992  
 #description Additional late promoter sequences of bacteriophage PI.  
 #accession S3855

SUMMARY #molecule-type DNA  
 #residues 26-126 #label SCH  
 #length 126 #molecular-weight 13588 #checksum 6874

Query Match Best Local Similarity 56.3%; Pred. No. 9.45e-02; Length 126;  
 Matches 9; Conservative 5; Mismatches 2; Indels 0; Gaps 0;

Db 76 ALNSALFLRRNGVQV 91  
 Qy 5 TINGQOLLYFRANGVQV 20

RESULT 14  
 ENTRY H70410 #type complete  
 TITLE UDP-N-acetylglucosamine 1-carboxyvinyltransferase - Aquifex aeolicus  
 ORGANISM #formal\_name Aquifex aeolicus  
 DATE 08-May-1998 #sequence\_revision 08-May-1998 #text\_change  
 13-Sep-1998

ACCESSIONS H70410  
 REFERENCE #authors Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Overbeek, R.; Snead, M.A.; Kellie, M.; Aufay, M.; Huber, R.; Feldman, R.A.; Short, J.M.; Olson, G.J.; Swanson, R.V.  
 #journal Nature (1998) 392:353-358  
 #title The complete genome of the hyperthermophilic bacterium

RESULT 15  
 ENTRY T00323 #type complete  
 TITLE chitinase (EC 3.2.1.14) B - Clostridium paraputreficum  
 ORGANISM #formal\_name Clostridium paraputreficum  
 DATE #length 425 #molecular-weight 47259 #checksum 7026

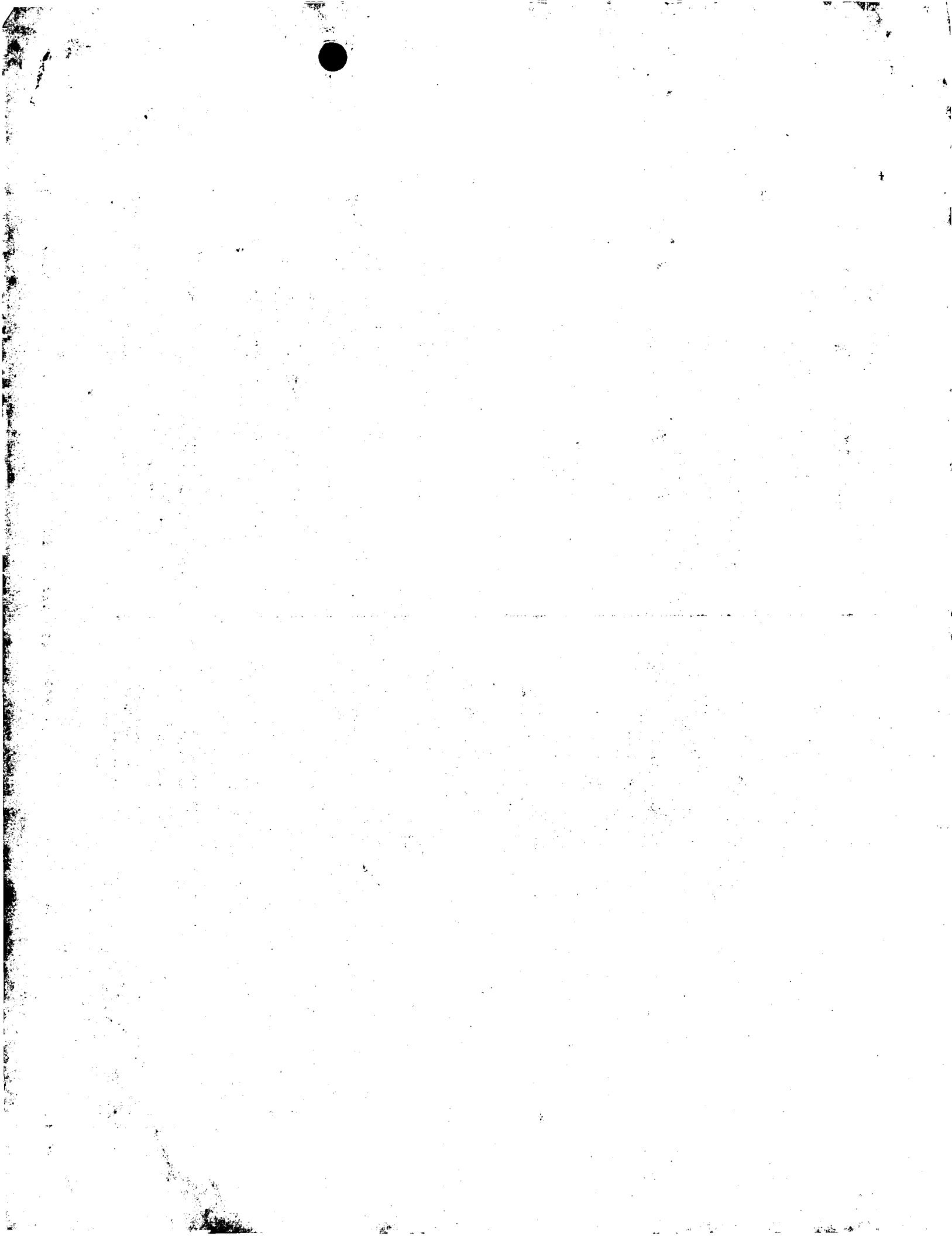
ACCESSIONS T00323  
 REFERENCE #authors Ohmiya, K.; Sakka, K.; Kimura, T.; Karita, S.; Morimoto, K.  
 #journal J. Bacteriol. (1997) 179:7306-7314  
 #description Cloning, sequencing, and expression of the gene encoding Clostridium paraputreficum chitinase ChiB and analysis of the functions of novel cadherin-like domains and a chitin-binding domain.

SUMMARY #accession T00323  
 #status Preliminary  
 #residues 1-831 #label OHM  
 #cross-references EMBL:AB001874; NID:d1175645; PID:di024701  
 #length 831 #molecular-weight 90021 #checksum 7845

Query Match Best Local Similarity 36.4%; Pred. No. 5.03e-01; Length 831;  
 Matches 8; Conservative 9; Mismatches 5; Indels 0; Gaps 0;

Db 525 TAATKVKNMKYIKTNGVATG 546  
 Qy 1 NGARTINGQOLLYFRANGVQVKG 22

Search completed: Tue Jan 11 15:47:33 2000  
 Job time : 28 secs.





ALTERNATE\_NAMES sucrose 6-glucosyltransferase  
 ORGANISM #formal\_name Streptococcus mutans  
 DATE 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change  
 22-Nov-1996  
 ACCESSIONS JT0345; C33135  
 REFERENCE Ueda, S.; Shiroza, T.; Kuramitsu, H.K.  
 Gene (1988) 69:101-109  
 Sequence analysis of the gtfB gene from *Streptococcus mutans*  
 GS-5;  
 #cross-references MUID:89137980  
 #accession JT0345  
 #molecule\_type DNA  
 #residues 1-1375 #label UED  
 #experimental\_source GS-5  
 REFERENCE A33135  
 #authors J. Bacteriol. (1987) 169:4263-4270  
 #journal title Sequence analysis of the gtfB gene from *Streptococcus mutans*.  
 #cross-references MUID:87308013  
 #accession C33135  
 #status preliminary  
 #residues 1-349 #label SHI  
 #molecule\_type DNA  
 #cross-references GB:M17361  
 GENETICS  
 #gene gtfC  
 FUNCTION #description catalyzes the synthesis of both water-soluble and  
 #superfamily cpl repeat homology  
 #domain cpl repeat homology #label CP2\\  
 CLASSIFICATION #water-insoluble glucans from glucose.  
 #domain cpl repeat homology #label CP3  
 KEYWORDS duplication; glycosyltransferase; hexosyltransferase;  
 transferase  
 FEATURE 1-34  
 #domain signal sequence #status predicted #predicted #label SIG\\  
 35-1375  
 #domain cpl repeat homology #label CP1\\  
 #domain cpl repeat homology #label CP2\\  
 #domain cpl repeat homology #label CP3  
 SUMMARY #length 1375 #molecular-weight 153021 #checksum 7015  
 Query Match 91.9%; Score 136; DB 2; Length 1375;  
 Best Local Similarity 95.2%; Pred. No. 3; 1.38e-14;  
 Matches 20; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Db 507 ANDHLSILEAWSNDTPYLHD 527  
 1 ANDHLSILEAWSNDTPYLHD 21  
 RESULT 3  
 ENTRY A38175  
 TITLE #type complete  
 ORGANISM #formal\_name *Streptococcus sobrinus*  
 DATE 28-Aug-1992 #sequence\_revision 28-Aug-1992 #text\_change  
 09-Sep-1997  
 ACCESSIONS A38175  
 REFERENCE #authors K.; Nagawa, H.; Kodama, T.; Ohta, H.; Fukui, K.; Kato,  
 J. Bacteriol. (1991) 173:989-996  
 Peptides sequences for sucrose splitting and glucan binding  
 Within *Streptococcus sobrinus* glucosyltransferase  
 (water-insoluble glucan synthetase).  
 #cross-references MUID:91123227  
 #accession A38175  
 #status preliminary  
 #molecule\_type DNA  
 #residues 1-1592 #label ABO  
 #cross-references GB:D90213; MUID:g217032; PID:d1014946; PID:g217033  
 CLASSIFICATION #superfamily cpl repeat homology  
 KEYWORDS glycosyltransferase; hexosyltransferase  
 FEATURE

RESULT 4  
 ENTRY A41483  
 TITLE #type complete  
 #cross-references (EC 2.4.1.-) gtfS precursor -  
 ORGANISM #formal\_name *Streptococcus sobrinus*  
 DATE 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change  
 09-Sep-1997  
 ACCESSIONS A41483  
 REFERENCE #authors Gilmore, R.S.; Russell, R.B.; Ferretti, J.J.  
 #journal title Infect. Immun. (1990) 58:2455-2458  
 #description Analysis of the *Streptococcus downei* gtfS gene, which  
 specifies a glucosyltransferase that synthesizes soluble  
 glucans.  
 #cross-references MUID:90316665  
 #accession A41483  
 #molecule\_type DNA  
 #residues 1-1365 #label GIL  
 #cross-references GB:M30943; MUID:9153652; PID:g153653  
 GENETICS  
 #gene gtfS  
 CLASSIFICATION #superfamily cpl repeat homology  
 #domain cpl repeat homology  
 #domain cpl repeat homology #label CP1\\  
 KEYWORDS glycosyltransferase; hexosyltransferase  
 SUMMARY #length 1365 #molecular-weight 151590 #checksum 823  
 Query Match 69.9%; Score 102; DB 2; Length 1365;  
 Best Local Similarity 71.4%; Pred. No. 4.40e-07;  
 Matches 15; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 Db 467 ANDHLSILEAWSNDTPYLHD 487  
 Qy 1 ANDHLSILEAWSNDTPYLHD 21  
 RESULT 5  
 ENTRY A45866  
 TITLE #type complete  
 ORGANISM #formal\_name *Streptococcus mutans*  
 DATE 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change  
 07-Sep-1995  
 ACCESSIONS A45866  
 REFERENCE #authors Honda, O.; Kato, C.; Kuramitsu, H.K.  
 #journal J. Gen. Microbiol. (1990) 136:2093-2105  
 #title Nucleotide sequence of the *Streptococcus mutans* gtfD gene  
 #description encoding the glucosyltransferase-S enzyme.  
 #accession A45866  
 #status preliminary  
 #molecule\_type DNA  
 #residues 1-1431 #label HON  
 #cross-references GB:M29396  
 CLASSIFICATION #superfamily cpl repeat homology  
 KEYWORDS glycosyltransferase; hexosyltransferase  
 FEATURE

181-201  
127-146  
1192-1211  
1237-1276  
1277-1297  
1321-1340  
1341-1361  
1355-1404

**SUMMARY**

Query Match  
Best Local Similarity 68.2%; Score 101; DB 2; Length 1431;  
Matches 14; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

**RESULT** 6

ENTRY S22737 #type complete  
ORGANISM glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius  
DATE 09-Sep-1997

**ACCESSIONS**  
REFERENCE S22737; S28810; BA4811; S22727  
#authors Jacques, N.  
submitted to the EMBL Data Library, March 1992  
#accession S22737

#molecule\_type DNA  
##residues 1-1599 ##label JAC  
##cross-references EMBL:211872; NID:947530; PID:947531

**REFERENCE**  
#authors A44811  
#journal Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.  
#title J. Gen. Microbiol. (1991) 137:2579-2593  
Molecular characterization of a cluster of at least two glucosyltransferase genes in Streptococcus salivarius ATCC 25975.

#cross-references MURD:92148377  
#accession S28810  
##molecule\_type DNA  
##cross-references EMBL:211873

**GENETICS**  
#gene gtfK  
#classification gtfK  
#KEYWORDS glycosyltransferase; hexosyltransferase

**SUMMARY**

Query Match  
Best Local Similarity 50.7%; Score 75; DB 2; Length 1518;  
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

**RESULT** 8

ENTRY B48445 #type complete  
TITLE glyceraldehyde-3-phosphate dehydrogenase (EC 1.2.1.12)  
ORGANISM Leishmania mexicana  
DATE 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 08-Sep-1997

**ACCESSIONS**  
REFERENCE A48445  
#authors Hannart, V.; Blaauw, M.; Kohl, L.; Allert, S.; Opperdoes, F.R.; Michels, P.A.M.  
#journal Mol. Biochem. Parasitol. (1992) 55:115-126  
#title Molecular analysis of the cytosolic and glycosomal glyceraldehyde-3-phosphate dehydrogenase in Leishmania mexicana.

#cross-references MURD:93065042  
#accession B48445  
##status preliminary  
##molecule\_type DNA  
##residues 1-331 ##label HAN  
##cross-references EMBL:IX65220; NID:9552; PID:9553

**CLASSIFICATION** #superfamily glyceraldehyde-3-phosphate dehydrogenase  
#KEYWORDS oxidoreductase

**SUMMARY**

Query Match  
Best Local Similarity 48.6%; Score 72; DB 2; Length 331;  
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

**RESULT** 9

ENTRY QQB840 #type complete  
TITLE BGLB2 protein - human herpesvirus 4 (strain B95-8)  
ORGANISM #formal\_name Human herpesvirus 4, Epstein-Barr virus  
DATE 25-Feb-1985 #sequence\_revision 25-Feb-1985 #text\_change 22-Jan-1999

**ACCESSIONS**  
REFERENCE C43044; JQ1381; A03784; A03794; S33036  
#authors Banerji, A.T.; Deininger, P.L.; Farrell, P.J.; Farrell, B.G.  
#journal Mol. Biol. (1983) 1:21-45  
#title Sequence analysis of the 17,166 bp EcoRI fragment C of B95-8  
Molecular characterization of a cluster of at least two glucosyltransferase genes in Streptococcus salivarius ATCC 25975.

#cross-references MURD:85035713  
#accession C43044  
##molecule\_type DNA



```

Db 1452 NKRLEQIKSWEETELSYLOE 1471 #title Trigamin. A low molecular weight peptide inhibiting
| : | : | : ||: fibrinogen interaction with platelet receptors expressed on
Qy 2 NDHLSILEAWSNDTPYLHD 21 #cross-references glycoprotein IIB-IIIA complex.
#accession MRID:88050981
#molecule_type protein

RESULT 13
ENTRY S58229 #type complete
TITLE pchA protein - Pseudomonas aeruginosa
ORGANISM #formal_name Pseudomonas aeruginosa
DATE 13-Mar-1998
ACCESSIONS 13-Mar-1998
REFERENCE S60203; S58229
SERINO, L.; REIMMANN, C.; BAUR, H.; BEYELER, M.; VISCA, P.; HAAS, D.
Mol. Gen. Genet. (1995) 249:217-228
#journal Structural genes for salicylate biosynthesis from chortismate in Pseudomonas aeruginosa.
#cross-references MRID:9606539
#accession S60203
#status preliminary
#molecule_type DNA
#residues 1-476 #label SE2
#cross-references EMBL:XB2644; NCBI:Q9G124; PDB:9929780
CLASSIFICATION #superfamily isochorismate synthase
SUMMARY #length 476 #molecular-weight 52071 #checksum 7909
Query Match Best Local Similarity 47.1%; Score 64; DB 2; Length 476;
Matches 8; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
#note this protein is involved in salicylate biosynthesis
#cross-references EMBL:XB2644; NCBI:Q9G124; PDB:9929780
#superfamily isochorismate synthase
#length 476 #molecular-weight 52071 #checksum 7909
Db 44 DPLQVFGAWDROTPCL 60
Qy 3 DHLSILEAWSNDTPYL 19
RESULT 14
ENTRY A30055 #type complete
TITLE trigramin precursor - Indian green tree viper hemorrhagic proteinase (EC 3.4.24.); platelet aggregation inhibitor (disintegrin) trigramin
ORGANISM #formal_name Trimeresurus gramineus #common_name Indian green tree viper
DATE 16-Sep-1992 #sequence_revision 16-Sep-1992 #text_change
05-Sep-1997
ACCESSIONS S12589; A30065; A29784
REFERENCE Neper, M.P.; Jacobson, M.A.
#authors Journal
#title Sequence of a cDNA encoding the platelet aggregation inhibitor trigramin.
#cross-references MRID:90332429
#accession S12589
#molecule_type mRNA
#residues 1-480 #label NEE
#cross-references EMBL:X51530; NID:964407; PID:96408
#note translation of the signal sequence and the mature protein but not of the propeptide is given
A30065 Huang, T.F.; Holt, J.C.; Kirby, E.P.; Niewiarowski, S.
#authors #journal
#title Trigramin: primary structure and its inhibition of von Willebrand factor binding to glycoprotein Ib/IIa complex on human platelets.
#cross-references MRID:89229063
#accession A30065
#molecule_type protein
#residues 408-479 #label HUA
#cross-references E71447
#status preliminary; nucleic acid sequence not shown;
#molecule_type DNA
#residues 1-515 #label BEV
#cross-references MRID:98121113
#accession E71447
#status preliminary; nucleic acid sequence not shown;
#molecule_type DNA
#residues 1-515 #label BEV
#cross-references GB:297344; NID:92245126; PID:e327081; PID:92245133
#GENETICS
#map_Position 4COPP-4G3Q45
SUMMARY #length 515 #molecular-weight 55762 #checksum 303
Query Match Best Local Similarity 43.2%; Score 64; DB 2; Length 515;

```

Thu Jan 13 09:13:33 2000

US-09-290-049-1.rpr

Page 6

|    | Matches | 7;                  | Conservative | 7; | Mismatches | 5; | Indels | 0; | Gaps | 0; |
|----|---------|---------------------|--------------|----|------------|----|--------|----|------|----|
| Db | 128     | TEENLFILEAWRTIDRAYI | 146          |    |            |    |        |    |      |    |
| Qy | 1       | ANDHLSILEAWSNDTPYL  | 19           |    |            |    | :      |    |      |    |

Search completed: Tue Jan 11 15:37:09 2000  
Job time : 14 secs.



**RESULT** 2  
**ID** R43696 standard; peptide; 22 AA.  
**AC** R43696;  
**DT** 20-MAY-1994 (first entry)  
**DE** GT subsequence.  
**KW** glucosyltransferase; vaccine; T-cell; B-cell; reaction;  
**KW** immunoresponse; peptidyl core matrix; dental caries; diphtheria;  
**KW** tetany; measles; polo.  
**OS** synthetic.  
**PN** WO9322341-A.  
**IL** -NOV-1993.  
**PR** 30-APR-1993; U04094.  
**PR** 01-MAY-1992; US 877295.  
**PA** (FORS-) FORSYTH DENTAL INFIRMARY FOR CHILDREN.  
**PT** Smith DJ, Taubman MA;  
**DR** WPI; 93-369721/46.  
**PT** synthetic glucosyl-transferase peptide(s) - provoke T- and B-cell reactions in mammals, and are effective against dental caries.  
**PS** Claim 3; Page 23; 28PP; English.  
**CC** The sequences (R43694-98) are subsequences from GT, they provoke T- and B-cell responses in mammals. The can be used to produce a vaccine comprising 2 of these sequences attached to a peptidyl core matrix. They are useful in producing T-cell responses to the GT enzymes causing a reduction of colonisation or accumulation of mutans streptococcal strains in preventing dental caries.  
**CC** CC vaccine comprising 2 of these sequences attached to a peptidyl core matrix. They are useful in producing T-cell responses to the GT enzymes causing a reduction of colonisation or accumulation of mutans streptococcal strains in preventing dental caries.  
**CC** CC vaccine comprising 2 of these sequences attached to a peptidyl core matrix. They are useful in producing T-cell responses to the GT enzymes causing a reduction of colonisation or accumulation of mutans streptococcal strains in preventing dental caries.  
**SQ** Sequence 22 AA;

**Query Match** 73.5%; Score 114; DB 8; Length 22;  
**Best Local Similarity** 77.3%; **Pred.** No. 4.41e-05;  
**Matches** 17; **Conservative** 3; **Mismatches** 2; **Indels** 0; **Gaps** 0;

**Db** 1 tgaqtikgqklyfkangqvg 22  
**Qy** 1 tGARTINGOLIYFRANGVQKG 22

**RESULT** 3  
**ID** W34164 standard; peptide; 24 AA.  
**AC** W34164;  
**DT** 18-FEB-1998 (first entry)  
**DE** GTF antigenic peptide #4 linked to polylysine core.  
**KW** Glucosyltransferase; GTF; enzyme; immunogen; catalytic domain; vaccine;  
**KW** surface domain; glucan-binding domain; mutans streptococcal strain;  
**KW** immune response; glucan-binding activity; dental caries prevention;  
**OS** synthetic.  
**OS** Streptococcus mutans.  
**FH** Location/Qualifiers  
**FT** Peptide 1..22  
**FT** /label= "GTF antigenic peptide #4 (see W34159)"  
**FT** /note= "attached to the dendritic polylysine core via the alpha-amino group of Lys(23); a second copy of the antigenic 22mer is linked to Lys(23) via the omega amino group"  
**FT** Modified\_site 23 /note= "Lys(23) is linked to one copy of the antigenic peptide through the alpha-amino group, and to a second copy of the peptide (not shown) via the omega amino group".  
**FT** Modified\_site 24 /note= "the alpha amino acid group of Lys(24) forms a peptide linkage with the carboxyl amino group of Lys(23); the omega amino group of Lys(24) forms a peptide bond with a second Lys residue analogous to Lys(23)"  
**FT**

**RESULT** 4  
**ID** R32925 standard; Protein; 1592 AA.  
**AC** R32925;  
**DT** 28-JUN-1993 (first entry)  
**DE** Glucosyltransferase I.  
**KW** GT-1; Streptococcus; dental; caries.  
**KW** Streptococcus sobrinus.  
**OS** J05023188-A.  
**PD** 02-FEB-1993;  
**PR** 25-JUL-1991; JP-186592.  
**PA** (FUKU/) FUKU I.  
**PA** (KATO/) KATO K.  
**DR** WPI; 93-079449/0.  
**DR** N-PSDB: 037760.  
**PT** DNA sequence glucosyl:transferase-I - comprises Streptococcus sobrinus DNA sequence with at least one nucleotide added or deleted  
**PS** Claim 13; Page 15; 28PP; Japanese.  
**CC** The DNA sequence from Streptococcus sobrinus strain 6715 encodes glucosyltransferase-I (and mutants). The DNA was obtd. by treating S. sobrinus 6715 with mutanolisin, extracting the chromosomal DNA, partially digesting with S. sobrinus 6715 with mutanolisin, extracting the chromosomal DNA, transforming it. A pET-109 (E. coli) vector was isolated and sequenced. The clone may be used in the development of a drug for dental caries.  
**SQ** Sequence 1592 AA;

**Query Match** 72.3%; Score 112; DB 6; Length 1592;  
**Best Local Similarity** 77.3%; **Pred.** No. 7.62e-05;  
**Matches** 17; **Conservative** 2; **Mismatches** 3; **Indels** 0; **Gaps** 0;

**Db** 1298 tgaqtikgqklyfkangqvg 1319

|                       |   |                            |  |
|-----------------------|---|----------------------------|--|
| Qy                    | 1   | TGARTINGOLLYFRANGVQVKG 22  |  |
| PT                    | C. difficile type toxins, and to treat C. difficile intoxication,   |                            |  |
| PR                    | partic. diarrhoea   |                            |  |
| PS                    | partic. diarrhoea   |                            |  |
| CC                    | Claim 53; Page 310-312; 434PP; English.   |                            |  |
| CC                    | A recombinant protein (R95014) comprises amino acids 1870-2680 of Clostridium difficile VPI strain 10463 toxin A (see also R95016), by a cytotoxin associated with diarrhoeal disease. It was optd. by PCR amplification of genomic DNA, and was expressed as a soluble fusion protein, with maltose binding protein as fusion partner, in transformed E. coli host cells. The soluble fusion protein can be used to raise avian antibodies useful as antitoxins or diagnostics.  |                            |  |
| CC                    | Sequence 811 AA;  |                            |  |
| OS                    | N-PADB; T13139.   |                            |  |
| PT                    | plants contg. new bacterial DNA encoding glucosyl transferase activity - retain higher levels of stored carbohydrate(s) in a form readily digestible by ruminants   |                            |  |
| PT                    | Claim 4; Page 16-20; 31PP; English.   |                            |  |
| CC                    | The sequence represents an alpha-D-glucosyltransferase from Streptococcus salivarius. The enzyme is primer-independent, and produces soluble glucan from sucrose. A gene encoding the enzyme may be cloned and expressed in Escherichia coli using a subclone of phage lambda-C13, e.g. plasmid PGSG501 or plasmid PGSG502. The DNA may also be expressed in a transgenic plant, to improve the level of stored carbohydrate in a pasture plant which normally contains low levels, or to prevent degradation of stored carbohydrate during plant senescence. Dextromay be isolated from the plant, for use as a food binder or pharmaceutical additive. Primer independence ensures that the enzyme will be functional in plants. The glucan is poorly degraded in plants but easily degraded by bacteria in the rumen sequence 1577 AA; |                            |  |
| Query Match           | 56.8%; Score 88; DB 16; Length 1577;  |                            |  |
| Best Local Similarity | 59.1%; Pred. No. 4 64e-02;  |                            |  |
| Matches               | 13; Conservative 3; Mismatches 6; Indels 0; Gaps 0;   |                            |  |
| Db                    | 1533 tgqgqngkqlyfdsgqrqkg 1554  |                            |  |
| Os                    | 1 TGARTINGOLLYFRANGVQVKG 22   |                            |  |
| RESULT                | 7   |                            |  |
| ID                    | R95017 standard; Protein: 812 AA.   |                            |  |
| AC                    | R95017;   |                            |  |
| DT                    | 08-JUL-1996 (first entry)   |                            |  |
| DE                    | C. difficile toxin A (aa1873-2684).   |                            |  |
| KW                    | Toxin A; cytotoxin; enterotoxin; fusion protein; antitoxin; vaccine;  |                            |  |
| CC                    | immunogen; diarrhoea; diagnosis; therapy;   |                            |  |
| OS                    | Clostridium difficile VPI strain 10463 (ATCC 10463).  |                            |  |
| PN                    | W09612802-A1.   |                            |  |
| PD                    | 02-MAY-1996.  |                            |  |
| PR                    | 23-OCT-1995; U13737.  |                            |  |
| PR                    | 24-OCT-1994; US-329154.   |                            |  |
| PR                    | 16-MAR-1995; US-405496.   |                            |  |
| PR                    | 14-APR-1995; US-422711.   |                            |  |
| PR                    | 07-JUN-1995; US-480004.   |                            |  |
| PA                    | (OPHI-) OPHTIAN PHARM INC.  |                            |  |
| PI                    | Firca JR, Kink JA, Padhye NV, Stafford DC, Thalley BS;  |                            |  |
| PI                    | Williams JA;  |                            |  |
| DR                    | WPI; 96-230603/23.  |                            |  |
| PT                    | Fusion proteins comprising non-toxin protein and part of toxin useful to form anti-toxins against Clostridium botulinum type A, and C. difficile type toxins, and to treat C. difficile intoxication,   |                            |  |
| PT                    | partic. diarrhoea   |                            |  |
| PS                    | Claim 79; Page 354-356; 434PP; English.   |                            |  |
| CC                    | A Protein (R95017) comprising amino acids 1873-2684 of Clostridium difficile VPI strain 10463 toxin A (see also R95016) was obtnd. by PCR amplification of genomic DNA. The PCR product can be expressed in transformed E. coli host cells as a soluble fusion protein, with maltose binding protein or a polyhistidine affinity tag as fusion partner. The resulting soluble fusion proteins, which are substantially endotoxin-free, can be used in anti-clostridial vaccines or to raise avian antibodies useful in clostridial antitoxin therapy for humans and animals.  |                            |  |
| SQ                    | Sequence 812 AA;  |                            |  |
| Query Match           | 43.2%; Score 67; DB 17; Length 812;   |                            |  |
| Best Local Similarity | 52.2%; Pred. No. 9 21e+00;  |                            |  |
| Matches               | 12; Conservative 3; Mismatches 7; Indels 1; Gaps 1;   |                            |  |
| Db                    | 588 tgirtidgkkyffntavatrg 610   |                            |  |
| Qy                    | 1 TGARTINGOLLYFRAN-GVQVKG 22  |                            |  |
| RESULT                | 8   |                            |  |
| ID                    | R95016 standard; Protein: 2710 AA.  |                            |  |
| AC                    | R95016;   |                            |  |
| DR                    | 08-JUL-1996 (first entry)   |                            |  |
| DE                    | C. difficile toxin A.   |                            |  |
| KW                    | Toxin A; cytotoxin; enterotoxin; fusion protein; antitoxin;   |                            |  |
| CC                    | vaccine; diarrhoea; diagnosis; therapy.   |                            |  |
| OS                    | Clostridium difficile VPI strain 10463 (ATCC 10463).  |                            |  |
| PT                    | Fusion proteins comprising non-toxin protein and part of toxin useful to form anti-toxins against Clostridium botulinum type A, and   |                            |  |
| PT                    | PR  | (OPHI-) OPHTIAN PHARM INC. |  |
| PI                    | Firca JR, Kink JA, Padhye NV, Stafford DC, Thalley BS;  |                            |  |
| PI                    | Williams JA;  |                            |  |
| DR                    | WPI; 96-230603/23.  |                            |  |

PN WO9612802-A1.  
 PD 02-MAY-1996.  
 PF 23-OCT-1995; U13737.  
 PR 24-OCT-1994; US-329154.  
 PR 16-MAR-1995; US-405495.  
 PR 14-APR-1995; US-422711.  
 PR 07-JUN-1995; US-480604.  
 PA (OPHI-) OPHIDIAN PHARM INC.  
 PI Williams JA; Padhye NV, Stafford DC, Thalley BS;  
 DR WPI; 95-23003/23.  
 N-PSDB; T2948.

Fusion proteins comprising non-toxin protein and part of toxin  
 product of the toxin A gene (T2948), a potent cytotoxin that  
 plays a direct role in damaging gastrointestinal tissues and is  
 associated with diarrhoeic disease. It can be obt'd. by expression in  
 transformed E. coli hosts of portions of DNA that together cover the  
 entire toxin A gene. Toxin A, and portions of it (see also  
 R95014-15 and R95017), pref. expressed as fusions to polyhistidine  
 affinity tags or maltose binding protein, are used to raise avian  
 antibodies useful as antitoxins or diagnostics, and in vaccine prodn.  
 Sequence 2710 AA;

Query Match 43.2%; Score 67; DB 17; Length 2710;  
 Best Local Similarity 52.2%; Pred. No. 9.21e+00;  
 Matches 12; Conservative 3; Mismatches 7; Indels 1; Gaps 1;

Db 2460 tqrlrtdgkkyffnttavatg 2482  
 QY 1 TGARTINGOLLYFRAN-GVQVK 22

RESULT 9  
 ID W68387 standard; Protein: 2710 AA.  
 AC W68387;  
 DT 07-DEC-1998 (first entry)  
 DE Clostridium difficile toxin A; toxin A; intoxication; immunogen;  
 KW Antitoxin; Vaccine; cytotoxin; toxin A; Intoxication; immunogen;  
 OS pseudomembranous enterocolitis.  
 KW Clostridium difficile.  
 PN WO9808540-A1.  
 PR 05-MAR-1998.  
 PR 28-AUG-1997; U15394.  
 PR 28-AUG-1996; US-704159.  
 PA (OPHI-) OPHIDIAN PHARM INC.  
 PI Thalley BS, Williams JA;  
 DR WPI; 95-23003/20.  
 N-PSDB; V30560.

Host cell containing recombinant expression vector encoding  
 Clostridium botulinum type B or E toxin - useful to treat humans  
 and other animals at risk of intoxication with clostridial toxin  
 Example 15; Page 220-230; 428PP; English.  
 This is the amino acid sequence of Clostridium difficile toxin A,  
 deduced from the coding region (see V30560) of the toxin A gene.  
 Toxin A is a potent cytotoxin that plays a direct role in damaging  
 gastrointestinal tissues. Severe cases of C. difficile  
 prevented by neutralising the effects of toxin A in the  
 gastrointestinal tract. Examples are provided of the production  
 of recombinant C. difficile toxin A in host cells and of the in  
 vivo neutralisation of toxin A by antibodies against recombinant  
 toxin A polypeptides. The invention specifically relates to  
 recombinant proteins derived from Clostridium botulinum toxins  
 (see WO8389-400) and their use as immunogens for the production of  
 vaccines and antitoxins.  
 Sequence 2710 AA;

Query Match 43.2%; Score 67; DB 35; Length 2710;  
 RESULT 10  
 ID R78621 standard; Protein: 404 AA.  
 AC R78621;  
 DT 11-JUN-1996 (first entry)  
 DE Chicken GaINAc-alpha-2,6-sialyltransferase P-B3.  
 KW GaINAc-alpha-2,6-sialyltransferase P-B3; anti-inflammatory;  
 viral infection; tumour; migration; inhibitor; sugar chain;  
 reagent; chicken.  
 OS Gallus gallus.  
 PN WO9518217-A1.  
 PR 06-JUL-1995.  
 PR 22-DEC-1994; J02182.  
 PR 24-DEC-1993; JP-34820.  
 PR 28-MAR-1994; JP-0915739.  
 PR 28-APR-1994; JP-091507.  
 PR (RIKA) INST PHYSICAL & CHEM RES.  
 PI Hamamoto T, Kojima N, Kurokawa N, Lee Y, Nakaoaka T;  
 PI Tsuji S;  
 DR WPI; 95-246383/32.  
 N-PSDB; 095325.  
 PR New GaINAc-alpha-2,6-sialyltransferase P-B1 and P-B3 - for e.g.  
 treatment of genetic disorders involving missing sugars  
 PT Claim 4; Pages 52-55; 70PP; Japanese.  
 PR 095325 encodes R78621 chicken GaINAc-alpha-2,6-sialyltransferase  
 P-B3. P-B3 can be used as a reagent for introducing human type  
 sugar chains onto proteins, or for the treatment of genetic  
 disorders involving missing sugar chains. It may also be used as  
 a tumour migration inhibitor, viral infection preventative and  
 CC as an anti-inflammatory.  
 Sequence 404 AA;

RESULT 11  
 ID W06416 standard; protein; 349 AA.  
 AC W06416;  
 DT 28-JAN-1997 (first entry)  
 DE Phosphotriesterase-related protein.  
 KW phosphotriesterase-related protein; mouse; PRP; recessive cpk; murine;  
 DE phosphotriesterase-related protein; mouse; PRP; recessive cpk; murine;  
 KW polycystic kidney disease; PKD; acute renal injury; autosomal dominant; kidney;  
 KW autosomal recessive; fluid-filled cyst; nephron; collecting duct; kidney;  
 KW renal parenchyma; renal failure; mammal; diagnostic marker; therapy;  
 KW acute organophosphate toxicity; phosphotriesterase inhibitor;  
 KW anti-cancer drug enhancer; anti-HIV drug enhancer.  
 OS Mus musculus.  
 PN US552313-A.  
 PR 03-SEP-1996.  
 PR 21-NOV-1994; 343027.  
 PR 21-NOV-1994; US-343027.  
 PA (UNIV ) UNIV KANSAS.  
 PI Calvet JP, Hou X, Magenheimer BS, Maser RL;  
 DR WPI; 96-412067/11.  
 N-PSDB; T43210.  
 PR Mouse phosphotriesterase-related protein DNA - used to develop  
 PT prods. for diagnosis of poly:cystic kidney disease or acute renal  
 PT failure.  
 PS Claim 1; Column 25-28; 23pp; English.  
 CC This sequence represents the mouse phosphotriesterase related protein

(mPRP). The mPRP sequence has abnormal underexpression in the recessive cpk murine model of a polycystic kidney disease (PKD), and has decreased expression following acute renal injury. PKD is a common disease (affecting 1 in 500-1000 individuals) and can be inherited as an autosomal dominant or as an autosomal recessive. PKD is characterised by the development of innumerable, large, fluid-filled cysts from the nephrons and collecting ducts of affected kidneys. It is thought that enlargement of the cysts interferes with functioning of the normal renal parenchyma, which eventually leads to renal failure. The mPRP, or other similar mammalian PRPs can also be used in therapeutic and protective treatments, such as for acute organophosphate toxicity. They can also be used to develop products to enhance the effectiveness of other types of therapy, e.g. phosphotriesterase inhibitors to enhance the effectiveness of certain anti-cancer, or anti-HIV drugs.

Sequence 349 AA;

Query Match

Best Local Similarity 40.0%; Score 62; DB 19; Length 349;  
Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 79 aireelyfkayg 91  
OY 5 TINGQIURRNG 17

RESULT 12  
ID W38549 standard; Protein; 207 AA.

AC W38549;  
DT 06-NOV-1998 (first entry)

DE S. pneumoniae peptide chain release factor 3 (RF-3).  
KW Streptococcus pneumoniae protein; genetic immunisation; antagonist; immunological response; inoculation; antibody production; inhibitor; KW cell immune response; antimicrobial compound; bacterial adhesion; KW extracellular matrix protein; protein-mediated cell invasion; wound; KW pathogenesis; S. pneumoniae.  
OS Streptococcus pneumoniae.

PN W0743303-A1.  
PD 20-NOV-1997.

PP 14-MAY-1997; U07950.

PR 14-MAY-1996; US-017-670.

PA (SMK ) SMITHKLINE BEECHAM CORP.

PA (SMK ) SMITHKLINE BEECHAM PLC.

PI Black MT, Hodgson JE, Knowles DJC, Nicholas RO,  
Stodola RK,  
DR WPI; 98-008793/01.

PR N-PSDB; T98605/01.

PT Novel Streptococcus pneumoniae peptide releasing factor polypeptide for diagnosing anti-microbial agents for treatment of bacterial

infections  
Claim 12; Page 327; 483pp; English.

This sequence represents a Streptococcus pneumoniae protein that, based on homology with *Bacteroides nodosus* protein, is a peptide chain release

factor 3 (RF-3), and is encoded by a DNA sequence of the invention.

The DNA sequences were isolated from *Streptococcus pneumoniae* strain NCIMB 40794. The *Streptococcus pneumoniae* protein of the

invention can be used to identify compounds which interact with and inhibit or activate the activity of the proteins. Antagonists can be used to treat diseases caused by *S. pneumoniae* proteins, through genetic

immunisation. They can also be used to induce an immunological response in a mammal by inoculation with the *S. pneumoniae* proteins or delivery of the encoding nucleic acids in a vector adequate to produce antibody and/or T cell immune responses to protect the animal from disease. The proteins can also be used to identify antimicrobial compounds which are capable of inhibiting their biactivity. In particular the proteins of the invention can be used to prevent adhesion of bacteria to mammalian

extracellular matrix proteins on in-dwelling devices or in wounds, to block protein-mediated mammalian cell invasion, and to block the normal progression of Pathogenesis in infections initiated other than by the

implantation of in-dwelling devices or other surgical techniques.  
Sequence 207 AA;

Query Match

Best Local Similarity 39.4%; Score 61; DB 39; Length 514;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 21 agkttiteqlyf 33  
OY 1 RQARTINGQOLYF 13

| Best Local Similarity | 61.5%;   | Pred. No.    | 3.85e+01;                                       |
|-----------------------|--|--------------|---|
| Matches               | 8;   | Conservative | 1;  |
| Mismatches            |  | Indels       | 0;  |
| Gaps                  |  | Peptides     | 0;  |
| Db                    | 21 agkttiteqlyf 33   | ID           | W83359  |
| AC                    |  | ID           | W83359;   |
| DT                    | 17-DEC-1996 (first entry)  | AC           | W83359;   |
| DE                    | Human cannabinoid GPR N-terminal sequence.   | DT           | 17-FEB-1999 (first entry)                       |
| KW                    | G-protein coupled receptor; ligand binding assay; transmembrane domain; schizophrenia; dopamine; cAMP; adenosine; thrombin; adrenoreceptor; opsin; muscarinic acetylcholine; endothelin; bombesin; endocrine; rhodopsin; | DE           | streptococcus pneumoniae prfC protein sequence. |

Query Match

Best Local Similarity 39.4%; Score 61; DB 34; Length 207;

QY 21 agkttiteqlyf 33  
OY 1 RQARTINGQOLYF 13

Query Match

Best Local Similarity 39.4%; Score 61; DB 39; Length 514;  
Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 21 agkttiteqlyf 33  
OY 1 RQARTINGQOLYF 13

KW odorant; cyromegalovirus; serotonergic.  
 OS Homo sapiens.  
 PN US5508384-A.  
 PD 16-APR-1996.  
 PF 10-SEP-1992; 943236.  
 PR 03-SEP-1993; US-118270.  
 PA (UINY ) UNIV NEW YORK STATE.  
 PI Murphy RB, Schuster DR;  
 DR WPI: 96-208785/21.

PT New dopamine receptor peptide - useful as antipsychotic agent, e.g. for treating schizophrenia.  
 PT Disclosure: Fig 8B(3); 184pp; English.  
 PS Proteins W02657-W02720 represent a range of G-protein coupled receptor (GPR) proteins selected from cAMP, adenosine, muscarinic acetylcholine, adrenergic, thrombin, endotoxin, bombesin, endocrine, rhodopsin, opsin, odorant, cyromegaloviral and other GPR proteins. The peptides W03578-W03651 represent the N-terminal fragments of the above proteins. The receptor proteins were used to design polypeptides, pref. based on the transmembrane domains, for use in G-protein coupled receptor ligand binding assays. The polypeptide fragments retain biological activity such as binding a GPR ligand or modulating GPR ligand binding to a GPR (see W02747-W02999 for examples of polypeptide fragments). The polypeptide fragments can be used in compositions for treating subjects suffering from a pathology related to a GPR abnormality e.g. a psychotic disorder such as schizophrenia.

Sequence 116 AA:  
 Query Match 38.7%; Score 60; DB 19; Length 116;  
 Best Local Similarity 50.0%; Pred. No. 4.8e+01;  
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 Qy 4 RTTIDDLVYGSNDQ 19  
 |  
 RESULT 15  
 ID R14196 standard; Protein; 472 AA.  
 AC R14196;  
 DT 19-DEC-1991 (first entry)  
 DE Human cannabinoid receptor.  
 KW Cannabis sativa; marijuana; drug test; substance K receptor.  
 OS Homo sapiens.  
 PN US7564075-A.  
 PD 03-SEP-1991.  
 PT 08-AUG-1990; US-554075.  
 PI (USSH ) NAT INST OF HEALTH.  
 PI Matsuda T, Brownstein M, Bonner T;  
 DR WPI: 91-30326/41.  
 DR N-PSDB; Q14003.

PT Human cannabinoid receptor.  
 PT Cannabis sativa; marijuana; drug test; substance K receptor.  
 PS Homo sapiens.  
 PT Human cannabinoid receptor - used for producing Disclosure; Fig 5; 25pp; English.  
 CC SK16 cDNA encoding the rat cannabinoid receptor was isolated from a rat cerebral cortex cDNA library. An EcoRI-XbaI fragment was used to screen a human cDNA library. A positive clone was identified and sequenced. There is ca. 97 per cent homology between this amino acid sequence deduced from the sequence of the positive human clone and rat cannabinoid receptor (see R14195). Recombinantly produced receptor can be used to screen for new drugs suitable for treatment of cannabinoid-treatable conditions, e.g. glaucoma, bronchial asthma, etc.  
 SQ Sequence 472 AA;

Query Match 38.7%; Score 60; DB 3; Length 472;  
 Best Local Similarity 50.0%; Pred. No. 4.8e+01;  
 Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

Db 14 rtttiddlvysndq 29  
 ||||:|||:||:||:  
 Qy 4 RTTIDDLVYGSNDQ 19

Search completed: Tue Jan 11 15:46:47 2000  
 Job time : 19 secs.



CC TOPOLOGY: linear  
 CC MOLECULE TYPE: peptide  
 CC SEQUENCE 22 AA; 2366 MW; 2475 CN;  
 SQ

Query Match 43.2%; Score 67; DB 2; Length 811;  
 Best Local Similarity 52.2%; Pred. No. 5.11e+00; Indels 3;  
 Matches 12; Conservative 3; Mismatches 7; Gaps 1;

Best Local Similarity 73.5%; Score 114; DB 1; Length 22;  
 Matches 17; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

Db 587 TGLRTIDGKKYIENNTAVATG 609  
 Qy |||:||| |||:||| |||:|||  
 1 TGARTINGQQLYFRANGVQVKG 22

RESULT 2 ID US-08-405-496A-7 STANDARD; PRT; 811 AA.  
 DE XX AC XXXXX  
 Sequence 7, Application US/08405496A

Sequence 7, Application US/08405496A  
 CC Patent No. 5915665  
 GENERAL INFORMATION:  
 APPLICANT: WILLIAMS, JAMES A.  
 TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM  
 NUMBER OF SEQUENCES: 30  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MEDLEN & CARROLL, LLP  
 STREET: 220 MONTGOMERY STREET, SUITE 2200  
 CITY: SAN FRANCISCO  
 STATE: CALIFORNIA  
 COUNTY: USA  
 ZIP: 94104

COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/405,496A  
 FILING DATE: 16-MAR-1995  
 CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/329,154  
 FILING DATE: 25-OCT-1994

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/161,907  
 FILING DATE: 02-DEC-1993

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/985,321  
 FILING DATE: 04-DEC-1992

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/429,791  
 FILING DATE: 31-OCT-1989

ATTORNEY/AGENT INFORMATION:  
 NAME: INGOLA, DIANE E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: OPHD-01308  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 705-8410  
 TELEFAX: (415) 397-8338  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 811 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 SEQUENCE 811 AA; 9121 MW; 3356332 CN;

RESULT 3 ID US-08-480-604A-7 STANDARD; PRT; 811 AA.  
 DE XX AC XXXXX  
 Sequence 7, Application US/08480604A  
 CC Patent No. 5736139  
 GENERAL INFORMATION:  
 APPLICANT: KINK, JOHN A.  
 APPLICANT: THALLEY, BRUCE S.  
 APPLICANT: PADHUA, NISHA V.  
 APPLICANT: FIRCA, JOSEPH R.  
 APPLICANT: STAFFORD, DOUGLAS C.  
 TITLE OF INVENTION: VACCINE AND ANTIOTOXIN FOR TREATMENT AND  
 NUMBER OF SEQUENCES: 32  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MEDLEN & CARROLL, LLP  
 STREET: 220 MONTGOMERY STREET, SUITE 2200  
 CITY: SAN FRANCISCO  
 STATE: CALIFORNIA  
 COUNTRY: UNITED STATES OF AMERICA  
 ZIP: 94104

COMPUTER READABLE FORM:  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/480,604A  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 424

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/422,711  
 FILING DATE: 14-APR-1995

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/405,496  
 FILING DATE: 16-MAR-1995

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/329,154  
 FILING DATE: 25-OCT-1994

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/161,907  
 FILING DATE: 02-DEC-1993

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/985,321  
 FILING DATE: 04-DEC-1992

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/429,791  
 FILING DATE: 31-OCT-1989

ATTORNEY/AGENT INFORMATION:  
 NAME: INGOLA, DIANE E.

REGISTRATION NUMBER: 40,027

REFERENCE/DOCKET NUMBER: OPHD-01763  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 705-8410  
 TELEFAX: (415) 397-8338  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 811 amino acids

CC TYPE: amino acid  
 CC STRANDEDNESS: unknown  
 CC TOPOLOGY: unknown  
 CC MOLECULE TYPE: Protein  
 CC SEQUENCE 811 AA; 91921 MW; 3356332 CN;

Db 587 TGLRTIDGKKYFNTAVAVTG 609  
 Qy ||| | :|| :| :| |  
 1 TGARTINGQOLLYFRAN-GVQVK 22

RESULT 4 ID US-08-480-604A-29 STANDARD; PRT; 812 AA.  
 XX AC xxxxxx

Sequence 29, Application US/08480604A  
 Sequence 29, Application US/08480604A  
 Patent No. 576139  
 GENERAL INFORMATION:  
 APPLICANT: KINK, JOHN A.  
 APPLICANT: THALLEY, BRUCE S.  
 APPLICANT: PADHIVE, NISHA V.  
 APPLICANT: FIRCA, JOSEPH R.  
 APPLICANT: STAFORD, DOUGLAS C.  
 TITLE OF INVENTION: VACCINE AND ANTIOTOXIN FOR TREATMENT AND  
 TITLE OF INVENTION: PREVENTION OF C. DIFFICILE DISEASE  
 NUMBER OF SEQUENCES: 32  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MEDLEN & CARROLL, LLP  
 STREET: 220 MONTGOMERY STREET, SUITE 2200  
 CITY: SAN FRANCISCO  
 STATE: CALIFORNIA  
 COUNTRY: UNITED STATES OF AMERICA  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/480,604A  
 FILING DATE: 07-JUN-1995  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/422,711  
 FILING DATE: 14-APR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/405,496  
 FILING DATE: 16-MAR-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/329,154  
 FILING DATE: 25-OCT-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/161,907  
 FILING DATE: 02-DEC-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/985,321  
 FILING DATE: 04-DEC-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/429,791  
 FILING DATE: 31-OCT-1989  
 ATTORNEY/AGENT INFORMATION:  
 NAME: INCOLIA, DIANE E.  
 REGISTRATION NUMBER: 40-027  
 REFERENCE/DOCKET NUMBER: OPHD-01763

CC TELECOMMUNICATION INFORMATION:  
 CC INFORMATION FOR SEQ ID NO: 29:  
 CC SEQUENCE CHARACTERISTICS:  
 LENGTH: 812 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 CC SEQUENCE 812 AA; 92022 MW; 3364670 CN;

Db 588 TGLRTIDGKKYFNTAVAVTG 610  
 Qy ||| | :|| :| :| |  
 1 TGARTINGQOLLYFRAN-GVQVK 22

RESULT 5 ID US-08-405-496A-6 STANDARD; PRT; 2710 AA.  
 XX AC xxxxxx

Sequence 6, Application US/08405496A  
 Sequence 6, Application US/08405496A  
 Patent No. 591665  
 GENERAL INFORMATION:  
 APPLICANT: WILLIAMS, JAMES A.  
 TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM  
 NUMBER OF SEQUENCES: 30  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: MEDLEN & CARROLL, LLP  
 STREET: 220 MONTGOMERY STREET, SUITE 2200  
 CITY: SAN FRANCISCO  
 STATE: CALIFORNIA  
 COUNTRY: USA  
 ZIP: 94104  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: FLOPPY disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30

CC CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/405,496A  
 FILING DATE: 16-MAR-1995  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/329,154  
 FILING DATE: 25-OCT-1994  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/161,907  
 FILING DATE: 02-DEC-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/985,321  
 FILING DATE: 04-DEC-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/429,791  
 FILING DATE: 31-OCT-1989  
 ATTORNEY/AGENT INFORMATION:  
 NAME: INCOLIA, DIANE E.  
 REGISTRATION NUMBER: 40-027  
 REFERENCE/DOCKET NUMBER: OPHD-01308  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 705-8410  
 TELEFAX: (415) 397-8338  
 INFORMATION FOR SEQ ID NO: 6:

CC SEQUENCE CHARACTERISTICS:  
 LENGTH: 2710 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 2710 AA; 308053 MW; 38468929 CN;  
 CC LENGTH: 2710 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC SEQUENCE 2710 AA; 308053 MW; 38468929 CN;  
 Query Match 43.2%; Score 67; DB 2; Length 2710;  
 Best Local Similarity 52.2%; Pred. No. 5.11e+00;  
 Matches 12; Conservative 3; Mismatches 7; Indels 1; Gaps 1;  
 ID US-08-480-604A-6 STANDARD; PRT; 2710 AA.  
 XX DE Sequence 6, Application US/08480604A  
 XX DE Sequence 6, Application US/08480604A  
 CC Patent No. 573639  
 CC GENERAL INFORMATION:  
 CC APPLICANT: KINK, JOHN A.  
 CC APPLICANT: THALLEY, BRUCE S.  
 CC APPLICANT: PADHYE, NISHA V.  
 CC APPLICANT: FIRCA, JOSEPH R.  
 CC APPLICANT: STAFORD, DOUGLAS C.  
 CC TITLE OF INVENTION: VACCINE AND ANTIOTOXIN FOR TREATMENT AND  
 CC NUMBER OF SEQUENCES: 32  
 CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: MEDLEN & CARROLL, LLP  
 CC STREET: 220 MONTGOMERY STREET, SUITE 2200  
 CC CITY: SAN FRANCISCO  
 CC STATE: CALIFORNIA  
 CC COUNTRY: UNITED STATES OF AMERICA  
 CC ZIP: 94104  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/480,604A  
 CC FILING DATE: 07-JUN-1995  
 CC CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 08/422,711  
 CC FILING DATE: 14-APR-1995  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 08/405,496  
 CC FILING DATE: 16-MAR-1995  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 08/329,154  
 CC FILING DATE: 25-OCT-1994  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 08/161,907  
 CC FILING DATE: 02-DEC-1993  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 07/985,321  
 CC FILING DATE: 04-DEC-1992  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 07/429,791  
 CC FILING DATE: 31-OCT-1989  
 CC ATTORNEY/AGENT INFORMATION:  
 NAME: INGOLA, DIANE E.  
 REGISTRATION NUMBER: 46,027  
 XX DT Sequence 7, Application US/08666367B  
 XX DT Sequence 7, Application US/08666367B  
 CC Patent No. 5851042  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Shuichi TSUJI et al.  
 CC TITLE OF INVENTION: NOVEL SUGAR-CHAIN SYNTHETASE AND PROCESS FOR  
 CC NUMBER OF SEQUENCES: 8  
 CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Wenderoth, Lind & Ponack  
 CC STREET: 805 Fifteenth Street, N.W., #700  
 CC CITY: Washington  
 CC STATE: D.C.  
 CC COUNTRY: U.S.A.  
 CC ZIP: 20005  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.5 inch, 1.44 mb  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: MS-DOS  
 SOFTWARE: Wordperfect 5.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/666,367B  
 FILING DATE: August 19, 1996  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/161,907  
 FILING DATE: 02-DEC-1993  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/985,321  
 FILING DATE: 04-DEC-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/429,791  
 FILING DATE: 31-OCT-1989  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Warren M. Cheek, Jr.  
 REGISTRATION NUMBER: 33,367  
 REFERENCE/DOCKET NUMBER:  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-371-8850  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 404 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 ORIGINAL SOURCE:  
 ORGANISM: G. gallus (chicken)  
 SEQUENCE 404 AA; 45826 MW; 844143 CN;

Query Match 40.6%; Score 63; DB 2; Length 404;  
 Best Local Similarity 45.0%; Pred. No. 1.32e-01; Length 514;  
 Matches 9; Conservative 5; Mismatches 6; Indels 0; Gaps 0;

Db 196 GRAIDAHDLVFRNGAITKG 215  
 Qy 3 ARTINGQOLYFRANGVQVKG 22

RESULT 8 STANDARD; PRT; 349 AA.

Sequence 12, Application US/08343027A

Patient No. 5552313

GENERAL INFORMATION:

APPLICANT: Calvet, James P., Xiaoying Hou, Robin L. Maser and Brenda S. Magen, TITLE OF INVENTION: DNA Encoding Mouse Phosphotriesterase-Related Protein NUMBER OF SEQUENCES: 12

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dr. Benjamin A. Adler, GILBRETH & ADLER, P.C.  
 STREET: 8011 Candle Lane  
 CITY: Houston  
 STATE: Texas  
 COUNTRY: USA  
 ZIP: 77071

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
 COMPUTER: Apple Macintosh  
 OPERATING SYSTEM: Macintosh  
 SOFTWARE: Microsoft Word for Macintosh

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/343,027A  
 FILING DATE:  
 CLASSIFICATION: 435  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Adler, Dr. Benjamin A.  
 REGISTRATION NUMBER: 35,423  
 REFERENCE/DOCKET NUMBER: D5728

TELECOMMUNICATION INFORMATION:

TELEPHONE: 713-777-2311  
 TELEFAX: 713-777-6908

INFORMATION FOR SEQ ID NO: 12:

SEQUENCE CHARACTERISTICS:

LENGTH: 349  
 TYPE: amino acid  
 TOPOLOGY: Linear  
 MOLECULE TYPE: Protein  
 HYPOTHETICAL: No  
 ANTI-SENSE: No  
 ORIGINAL SOURCE:  
 STRAIN:  
 INDIVIDUAL ISOLATE:  
 DEVELOPMENTAL STAGE:  
 TISSUE TYPE:  
 CELL TYPE:  
 CELL LINE:  
 SEQUENCE 349 AA; 39253 MW; 609827 CN;

Query Match 40.0%; Score 62; DB 1; Length 349;  
 Best Local Similarity 53.0%; Pred. No. 1.67e-01; Length 514;  
 Matches 7; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Db 79 AIRFELLYKAYG 91  
 Qy 5 TINGQOLYFRANG 17

RESULT 9 STANDARD; PRT; 514 AA.

Sequence 2, Application US/08865311

Patient No. 5019664

GENERAL INFORMATION:

APPLICANT: Pearson, Stewart C.  
 TITLE OF INVENTION: No. 5919664el Compounds  
 NUMBER OF SEQUENCES: 2

CORRESPONDENCE ADDRESS:

ADDRESSEE: Dechert, Price & Rhoads  
 STREET: 4000 Bell Atlantic Tower, 1717 Arch Stree  
 CITY: Philadelphia  
 STATE: PA  
 COUNTRY: USA  
 ZIP: 19103-2793

COMPUTER READABLE FORM:

MEDIUM TYPE: Discrete  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: Windows 95  
 SOFTWARE: FASTSEQ for Windows Version 2.0b

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/865,311  
 FILING DATE: 29-MAY-1997  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Falk, Stephen T  
 REGISTRATION NUMBER: 36,795  
 REFERENCE/DOCKET NUMBER: GM10004  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 215-994-2488  
 TELEFAX: 215-994-2222  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 514 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 CC TOPOLOGY: linear  
 SQ SEQUENCE 514 AA; 58470 MW; 1375700 CN;

Query Match 39.4%; Score 61; DB 2; Length 514;  
 Best Local Similarity 61.5%; Pred. No. 2.11e-01; Length 514;  
 Matches 8; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Db 21 AGKTTITEQLYF 33  
 Qy 1 TGARTINGQOLYF 13

RESULT 10 STANDARD; PRT; 91 AA.

Sequence 8, Application US/08480604A

Sequence 8, Application US/08865311

Patent No. 5736139

GENERAL INFORMATION:

APPLICANT: Kink, John A.

DE Sequence 8, Application US/0840596A.  
 CC XX  
 Sequence 8, Application US/0840596A  
 CC Patent No. 5919665  
 CC GENERAL INFORMATION:  
 CC APPLICANT: WILLIAMS, JAMES A.  
 CC TITLE OF INVENTION: VACCINE FOR CLOSTRIDIUM BOTULINUM  
 CC NUMBER OF SEQUENCES: 30  
 CC TITLE OF INVENTION: NEUROTOXIN  
 CC NUMBER OF SEQUENCES: 30  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: MEDLEN & CARROLL, LLP  
 CC STREET: 220 MONTGOMERY STREET, SUITE 2200  
 CC CITY: SAN FRANCISCO  
 CC STATE: CALIFORNIA  
 CC COUNTRY: UNITED STATES OF AMERICA  
 CC ZIP: 94104  
 COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: FLOPPY disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patentin Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/480,604A  
 CC FILING DATE: 07-JUN-1995  
 CC CLASSIFICATION: 424  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 08/422,711  
 CC FILING DATE: 14-APR-1995  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 08/405,496  
 CC FILING DATE: 16-MAR-1995  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 08/329,154  
 CC FILING DATE: 25-OCT-1994  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 08/161,907  
 CC FILING DATE: 02-DEC-1993  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 07/985,321  
 CC FILING DATE: 04-DEC-1992  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: US 07/429,791  
 CC FILING DATE: 31-OCT-1999  
 CC ATTORNEY/AGENT INFORMATION:  
 NAME: INGOLIA, DIANE E.  
 REGISTRATION NUMBER: 40,027  
 REFERENCE/DOCKET NUMBER: OPHD-01763  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 705-8410  
 FAX: (415) 397-8338  
 TELEFAX: (415) 397-8338  
 CC INFORMATION FOR SEQ ID NO: 8:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 91 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: unknown  
 TOPOLOGY: unknown  
 MOLECULE TYPE: protein  
 SEQUENCE 91 AA; 10432 MW; 39683 CN;  
 Query Match 38.1%; Score 59; DB 1; Length 91;  
 Best Local Similarity 46.7%; Pred. No. 3.36e+01;  
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
 Db 3 KIINGKHFVNNDGV 17  
 : ||| :||| :|||  
 Qy 4 RTINGQOLLYFRANGV 18  
 RESULT 11  
 ID US-08-405-496A-8 STANDARD PRT: 91 AA.  
 XX  
 AC XXXXXX  
 DT  
 XX  
 RESULT 12  
 ID US-08-485-607-6 STANDARD PRT: 453 AA.  
 XX  
 AC XXXXXX  
 DT  
 XX  
 Sequence 6, Application US/08485607  
 CC Sequence 6, Application US/08485607  
 CC Patent No. 5792627

CC GENERAL INFORMATION:  
 CC APPLICANT: Robert Haselkorn and Piotr Gornicki  
 CC TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA  
 CC NUMBER OF SEQUENCES: 116  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Arnold, White & Durkee  
 CC STREET: 321 NO. 579627th Clark Street  
 CC CITY: Chicago  
 CC STATE: Illinois  
 CC COUNTRY: USA  
 CC ZIP: 60610

CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy Disk  
 CC COMPUTER: IBM PC Compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/422,560A  
 CC FILING DATE: 14-APR-1995  
 CC CLASSIFICATION: 800  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: 07/956,700  
 CC FILING DATE: 07-JUN-1995  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Thomas E. No. 5792627thrup  
 CC REGISTRATION NUMBER: 33,268  
 CC REFERENCE/DOCKET NUMBER: ARCD:058  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 1-312-744-0090  
 CC TELEX/FAX: 1-312-755-4489  
 CC INFORMATION FOR SEQ ID NO: 6:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 453 amino acids  
 CC TYPE: Amino acid  
 CC STRANDEDNESS: Single  
 CC TOPOLOGY: Linear  
 CC MOLECULE TYPE: Peptide  
 CC SEQUENCE: 453 AA; 49608 MW; 983857 CN;  
 SQ

Query Match 38.1%; Score 59; DB 2; Length 453;  
 Best Local Similarity 36.8%; Pred. No. 3,36e+01;  
 Matches 7; Conservative 4; Mismatches 8; Indels 0; Gaps 0;

Db 315 RIAQGEALRROADIQLRG 333  
 Qy 4 RTINGQOLYFRANGVQVKG 22

CC RESULT 14  
 CC ID US-08-476-537-6 STANDARD; PRT; 453 AA.  
 CC XX  
 CC AC XXXXX  
 CC DT XXX  
 CC DB Sequence 6, Application US/08476537  
 CC Sequence 6, Application US/08476537  
 CC Patent No. 5756290  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Robert Haselkorn and Piotr Gornicki  
 CC TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA  
 CC NUMBER OF SEQUENCES: 116  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Arnold, White & Durkee  
 CC STREET: 321 NO. 5796290th Clark Street  
 CC CITY: Chicago  
 CC STATE: Illinois  
 CC COUNTRY: USA  
 CC ZIP: 60610

CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy Disk  
 CC COMPUTER: IBM PC Compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: ASCII-DOS  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/476,537  
 CC FILING DATE:  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: 07/956,700  
 CC FILING DATE: 10/21/92  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Thomas E. No. 5756290thrup

CC REGISTRATION NUMBER: 33\_268  
 CC REFERENCE/DOCKET NUMBER: ARCD-058  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: 1-312-744-0090  
 CC TELEFAX: 1-312-755-4489  
 CC INFORMATION FOR SEQ ID NO: 6:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 453 amino acids  
 CC TYPE: Amino acid  
 CC STRANDEDNESS: Single  
 CC TOPOLOGY: Linear  
 CC MOLECULE TYPE: Peptide

SQ SEQUENCE 453 AA; 49608 MW; 983857 CN;  
 Query Match 38.1%; Score 59; DB 2; Length 453;  
 Best Local Similarity 36.8%; Pred. No. 3.36e+01;  
 Matches 7; Conservative 4; Mismatches 8;  
 Indels 0; Gaps 0;

Qy 315 RIAQGEALRROADIQRG 333  
 4 RTINGQQLYFRANGVQVKG 22

Search completed: Tue Jan 11 15:49:31 2000  
 Job time : 7 secs.

|    | Matches | 7                   | Conservative | 4 | Mismatches | 8 | Indels | 0 | Gaps | 0 |
|----|---------|---------------------|--------------|---|------------|---|--------|---|------|---|
| Db | 315     | RIAQGEALRROADIQRG   | 333          |   |            |   |        |   |      |   |
| Qy | 4       | RTINGQQLYFRANGVQVKG | 22           |   |            |   |        |   |      |   |

Qy

RESULT 15  
 ID US-07-956-700B-6 STANDARD; PRN; 453 AA.

DE Sequence 6, Application US/07956700B

XX Sequence 6, Application US/07956700B  
 XX Patent No. 5539092

CC GENERAL INFORMATION:  
 CC APPLICANT: Robert Hagelkorn and Piotr Gornicki

CC TITLE OF INVENTION: Cyanobacterial and Plant Acetyl-CoA

CC NUMBER OF SEQUENCES: 116  
 CC CORRESPONDENCE ADDRESS:

CC ADDRESSEE: Arnold White & Durkee  
 STREET: 321 No. 5539092th Clark Street

CC CITY: Chicago  
 CC STATE: Illinois

CC COUNTY: USA  
 CC ZIP: 60610

CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy Disk

CC COMPUTER: IBM PC Compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS

CC SOFTWARE: ASCII-DOS  
 CC CURRENT APPLICATION DATA:

CC APPLICATION NUMBER: US/07/956,700B  
 CC FILING DATE: 1992/10/02

CC CLASSIFICATION: 800  
 CC ATTORNEY/AGENT INFORMATION:

CC NAME: Thomas E. NO. 5539092thrup  
 CC REGISTRATION NUMBER: 33\_268

CC REFERENCE/DOCKET NUMBER: ARCD-058  
 CC TELECOMMUNICATION INFORMATION:

CC TELEPHONE: 1-312-744-0090  
 CC TELEFAX: 1-312-755-4489

CC INFORMATION FOR SEQ ID NO: 6:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 453 amino acids

CC TYPE: Amino acid  
 CC STRANDEDNESS: Single  
 CC TOPOLOGY: Linear  
 CC MOLECULE TYPE: Peptide

SQ SEQUENCE 453 AA; 49608 MW; 983857 CN;  
 Query Match 38.1%; Score 59; DB 1; Length 453;  
 Best Local Similarity 36.8%; Pred. No. 3.36e+01;



AC R91047; DT 22-MAY-1996 (first entry)  
 DE Alpha-D-glucosyltransferase.  
 KW alpha-D-glucosyltransferase; primer-independent; soluble glucan;  
 sucrose; transgenic plant; cloning; Escherichia coli;  
 phage lambda-C13; vector; plasmid pGSG501; plasmid pGSG502;  
 gene transfer; crop improvement; storage carbohydrate; pasture;  
 feedstuff; senescence; dextran; binder; food; pharmaceutical.  
 KW Streptococcus salivarius strain ATCC 25975.  
 OS WO9606173-A1.

PD 29-FEB-1995; AU05277.  
 PR 24-AUG-1994; AU-007643.  
 PT (GIFF /) GIFFARD P M.  
 PA (JACQ /) JACQUES N A.  
 PA (SIMP /) SIMPSON C L.  
 PI Giffard PM, Jacques NA, Simpson CL;  
 DR WPI, 96-1376/15.  
 N-PSDB; TI3139.

Plants contg. new bacterial DNA encoding glucosyl transferase activity - retain higher levels of stored carbohydrate(s) in a form readily digestible by ruminants.

PT Claim 4; Page 16-20; 31pp; English.  
 CC The sequence represents an alpha-D-glucosyltransferase from Streptococcus salivarius. The enzyme is primer-independent, and produces soluble glucan from sucrose. A gene encoding the enzyme may be cloned and expressed in Escherichia coli using a subclone of phage lambda-C13, e.g. plasmid pGSG501 or plasmid pGSG502. The DNA may also be expressed in a transgenic plant, to improve the level of stored carbohydrate in a pasture plant which normally contains low levels, or to prevent degradation of stored carbohydrate from the plant, for use as a food binder or pharmaceutical additive. Primer independence ensures that the enzyme will be functional in plants. The glucan is poorly degraded in plants but easily degraded by bacteria in the rumen of grazing livestock.

SQ Sequence 1577 AA;

Query Match Best Local Similarity 52.6%; Score 71; DB 16; Length 1577; Matches 11; Conservative 5; Mismatches 2; Indels 1; Gaps 1;

Db 660 anyifvr-ahdievqavla 677  
 ::|:||| ||||| |::|  
 2 PSYFIRTAHDSEVQDIA 20

PT R22237 standard; Protein: 825 AA.  
 AC R22237;  
 DT 23-JUL-1992 (first entry)  
 DE Sequence of the "gb" homologue of ILTV encoded by the EcoRI "U" (unique) genome fragment.  
 DE Subunit vaccine; immunogen; glycoprotein; promoter.  
 OS Infectious laryngotracheitis virus.  
 FH Key peptide 1..16  
 FT /label= signal

PR 3

PT R22237 standard; Protein: 825 AA.  
 AC R22237;  
 DT 23-JUL-1992 (first entry)  
 DE Sequence of the "gb" homologue of ILTV encoded by the EcoRI "U" (unique) genome fragment.  
 DE Subunit vaccine; immunogen; glycoprotein; promoter.  
 OS Infectious laryngotracheitis virus.

FH Key peptide 1..16  
 FT /label= signal

PR 05-MAR-1992  
 PT WO9203554-A.  
 PR 23-AUG-1991; AU03831.  
 PR 24-AUG-1990; AU-001937.  
 PR (WEBB-) WEBSTER A PTY LTD.  
 PA Sheppard MG, Pridieux C, Johnson M, Fahey KJ, York JJ,  
 PI Kongsvan K;  
 DR WPI; 92-09698/12.  
 DR N-PSDB; Q22289.

Vaccines against ILTV for use in chickens - comprises non-infectious sub-units or live recombinant viruses and may be administered by aerosol.

PT Examined: Fig 11; 12pp; English.

CC The inventors claim a non-infectious subunit vaccine for use against ILTV which comprises a glycoprotein of ILTV. They also claim DNA

RESULT 5  
 ID W2812 standard; Protein: 344 AA.  
 AC W52812;  
 DT 07-JUL-1998 (first entry)  
 DE Human induced tumour protein.  
 KW Human induced tumour protein; HTP; treatment; prevention; cancer; melanoma; carcinoma.  
 OS Homo sapiens.  
 FH Location/Qualifiers

CC encoding the vaccine, a synthetic polypeptide displaying the antigenicity of the 205K complex or 60K ILTV glycoprotein and a recombinant ILTV with heterologous DNA inserted into a non-essential region of the genome; and a recombinant DNA mol., comprising an ILTV promoter region operatively linked to a heterologous DNA sequence or the ORF3 promoter region. KpnK/ORF3 is located 5' of the gp60 gene.

CC Sequence 825 AA;

Query Match Best Local Similarity 46.7%; Score 63; DB 4; Length 825; Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Db 570 ytfrvrgnaseeiv 584  
 ::|:| : ||:::::  
 Qy 4 PSYFIRTAHDSEVQDIA 18

RESULT 4  
 ID R79459 standard; Protein: 873 AA.  
 AC R79459;  
 DT 16-JAN-1996 (first entry)  
 DE Infectious laryngotracheitis virus gb glycoprotein.  
 KW vaccine; immunosay; detection; identification; avipox; chicken; fowl; probe; antisense; gene expression.  
 OS Gallid Herpesvirus 1.  
 FH Key msc\_difference 656..659  
 FT /note= "These amino acids are missing from the denoted polypeptide sequence given in the specification."  
 PT US5444831-A.  
 PD 22-AUG-1995.  
 PR 29-OCT-1991; 788123.  
 PR 19-NOV-1993; US-156866.  
 PA (URDE ) UNIV DELAWARE.  
 PT Keeler CL, Poulsen DJ;  
 DR WPI; 95-30091/39.  
 N-PSDB; 970351.

PT Isolated Infectious laryngotracheitis virus gb gene - used to develop prod. for detection and for vaccine(s), partic. for immunising fowl.

PT Claim 1; Columns 11-16; 1pp; English.

CC The infectious laryngotracheitis virus (ILTV) gb glycoprotein nucleic acids can be used as probes, for antisense control of gb gene expression or for production of gb polypeptides. The gb polypeptides can be used for the production of antibodies (both used in immunoassays) and in vaccines. Recombinant avipox virus expressing the ILTV gb can be used for immunising fowl, especially chickens.

SQ Sequence 873 AA;

Query Match Best Local Similarity 46.7%; Score 63; DB 14; Length 873; Matches 5; Conservative 7; Mismatches 3; Indels 0; Gaps 0;

Db 618 ytfrvrgnaseeiv 632  
 ::|:| : ||:::::  
 Qy 4 PSYFIRTAHDSEVQDIA 18

RESULT 5  
 ID W2812 standard; Protein: 344 AA.  
 AC W52812;  
 DT 07-JUL-1998 (first entry)  
 DE Human induced tumour protein.  
 KW Human induced tumour protein; HTP; treatment; prevention; cancer; melanoma; carcinoma.  
 OS Homo sapiens.  
 FH Location/Qualifiers



FT WO9515374-A1. /\*tag= a  
 PN 08-JUN-1995.  
 PD 30-NOV-1994; J02009.  
 PF 30-NOV-1993; JP-341061.  
 PR FUJI YAKUHIN KOGYO KK.  
 PA Sato H, Sasaki M, Shinagawa A;  
 PT WPI: 95-215255/28.  
 DR N-PSDB; Q92573.  
 PT Metallo-protease and monoclonal antibody recognising it - for  
 detection of cancer cells in medical diagnosis and research.  
 PS Claim 1; Pages 39-43; 67pp; Japanese.  
 CC 092573 encodes R75648 a human placenta derived metalloprotease. The  
 gene can be used as a probe for the detection of cancer cells, and  
 monoclonal antibodies specific for the metalloprotease can be used  
 for the treatment and diagnosis of cancer.  
 CC Sequence 582 AA;  
 SQ :|:||| :|: :|:  
 Qy 3 SYSFIR-TAHDSEVQDLI 19

RESULT 9  
 ID R32010 standard; Protein; 146 AA.  
 AC R32010; 1993 (first entry)  
 DE 15 kd antigen; cytolytic lymphocyte; infection; HIV; targetting.  
 KW Homo sapiens.  
 OS WO9301314-A.  
 PR 12-JUN-1992; US05117.  
 PR 19-FEB-1992; US-843949.  
 PA (DAND ) DANA FARMER CANCER INST INC.  
 PI Anderson RJ, Schlossman SF, Streuli M;  
 WPI; 93-045515/05.  
 DR N-PSDB; Q36555.  
 PT DNA encoding TIA-1 antigen - used for detecting and cytotoxic  
 killing of cytolytic lymphocyte(s) applicable for cancer  
 Disclosure: Page 37; 61pp; English.  
 PS RNA from cytolytic T cell clone T4Tc1, which expresses high levels  
 of TIA-1 antigen was used for the construction of a cDNA library in  
 lambda gt11. When this expression library was screened using TIA-1  
 monoclonal antibody, several bacteriophage expressing immunoreactive  
 fusion proteins were identified. Cross hybridisation analysis  
 identified three independent clones contg. the  
 largest cDNA insert was used to probe the original library for  
 isolation of full length cDNA encoding rpl3-TIA-1, a 15 KD TIA-1  
 antigen. The antigen is active in lymphocyte-mediated cytotoxicity  
 and can eliminate a wide variety of virus infected or transformed  
 target cells. TIA or TIA-1 antigen may be linked chemically or  
 recombinantly to cell targeting ligands such as growth factors,  
 hormone or antibodies and may be used to kill targetted cells, e.g.  
 cancer cells. See also R32009-17.  
 Sequence 146 AA;

Query Match 44.4%; Score 60; DB 6; Length 146;  
 Best Local Similarity 31.6%; Pred. No. 3.53e+01;  
 Matches 6; Conservative 8; Mismatches 4; Indels 1; Gaps 1;

Db 10 gysfvvfnshsaahivs 28  
 Qy :|:||| :|: :|: ::: 3 SYSFIR-TAHDSEVQDLI 20

RESULT 9  
 ID R12539 standard; Protein; 146 AA.  
 AC R12539; 1991 (first entry)  
 DE 24-SEP-1991 (first entry)  
 KW Cytotoxic T Lymphocyte-specific antigen.  
 CC CTL; natural killer cells; CD8; peripheral blood; TIA-1; immunogen.  
 PN EP-43400-A.  
 PR 31-JUL-1991.  
 PR 31-DEC-1990; 314456.  
 PR 05-JAN-1990; US-460678.  
 PA (DANA-) DANA FARMER CANCER.  
 PT Anderson RJ, Streuli M, Schlossman SF;  
 WPI; 91-202096/28.  
 DR N-PSDB; Q12525.  
 PT Monoclonal antibody to identify cytolytic lymphocytes - reactive  
 with a 15KD protein in cytolytic T lymphocytes and natural killer  
 cells.  
 PS Claim 10; Fig 1; 11pp; English.  
 CC This sequence corresponds to a 15KD immunogenic protein associated  
 with cytoplasmic granules in cytolytic T lymphocytes and natural  
 killer cells. The protein is found principally in a subpopulation of  
 CD8+ T lymphocytes from peripheral blood mononuclear cells. It  
 reacts with monoclonal antibody TIA-1 (HB 10319).  
 SQ Sequence 146 AA;

Query Match 44.4%; Score 60; DB 3; Length 146;  
 Best Local Similarity 31.6%; Pred. No. 3.53e+01;  
 Matches 6; Conservative 8; Mismatches 4; Indels 1; Gaps 1;

Db 10 gysfvvfnshsaahivs 28  
 Qy 3 SYSFIR-TAHDSEVQDLI 20

RESULT 11  
 ID R40840 standard; Protein; 156 AA.  
 AC R40840; 1994 (first entry)  
 DE 01-MAR-1994 (first entry)  
 KW Glialblastoma derived polypeptide.  
 CC Glialblastoma; differentiation; proliferation; growth; glia;  
 KW neuron; immunity; tumour; aplasia.  
 OS Homo sapiens.  
 FN Key  
 FT peptide 1.23  
 FT /label= signal\_peptide  
 FT protein 24..156  
 FT /label= Mature\_protein  
 PN EP-554428-A.  
 PR 08-SEP-1993.  
 PR 02-MAR-1993; 301571.  
 PR 03-MAR-1992; JP-081557.  
 PR 22-APR-1992; JP-129958.  
 PA (ONO ) ONO PHARM CO LTD.  
 PI Konishi M, Miyamoto T, Naitoh T;  
 DR WPI; 93-282547/36.  
 DR N-PSDB; Q47756.  
 PT New polypeptide from glioblastoma cell line - useful in treatment  
 of abnormal glial cells or neurons or for treating tumours  
 PS Claims 1-4; Page 14-15; 21pp; English.  
 CC The glioblastoma derived polypeptide is synthesised in and secreted  
 from a glioblastoma cell line containing this DNA fragment and  
 therefore may possess biological activities relating to  
 differentiation, proliferation and growth of glias or neurons,  
 relating to the function of immunity and relating to proliferation  
 and growth of tumours. The polypeptide may be useful for the  
 prevention of or in the treatment of aplasia or abnormal  
 proliferation of glias or neurons, depression or enhancement of  
 immunobiological activity. The DNA was amplified by PCR techniques  
 CC using primers (047754-55).  
 SQ Sequence 156 AA;

Query Match 44.4%; Score 60; DB 8; Length 156;  
 Best Local Similarity 42.9%; Pred. No. 3.53e+01;  
 Matches 9; Conservative 6; Mismatches 5; Indels 1; Gaps 1;

RESULT 10

**Db** 107 vpaldfrvsvphyevvia 127  
**ID** R32011; standard; Protein; 375 AA.  
**AC** ||| : ||| : ||| :|||  
**Qy** 1 VPSYS-FIRTAAHDSEVQDLIA 20

**RESULT** 12  
**ID** R32011; standard; Protein; 375 AA.  
**AC** R32011;  
**DT** 27-MAY-1993 (first entry)  
**DE** TIA1.  
**KW** Antigen; cytolytic lymphocyte; TIA-1 antigen related gene; infection; HIV; targetetting; cancer.  
**OS** Homo sapiens.  
**PN** WO9301314A.  
**PD** 21-JUN-1993; U05117.  
**PF** 12-JUL-1991; US-726607.  
**PR** 10-JUL-1991; US-726607.  
**PA** (DAND ) DANA FARMER CANCER INST INC.  
**PT** Anderson PJ, Schlossman SF, Streuli M;  
**DR** WPI; 93-045515/05.  
**N-PSDB:** 024556.

**PT** DNA encoding TIA-1 antigen - used for detecting and cytotoxic treatment of cytolytic lymphocyte(s) applicable for cancer

**PT** Disclosure: Page 39; 61PP; English.

**CC** Southern blots of genomic DNA probed with a TIA-1 antigen specific probe were unexpectedly complex suggesting that either the TIA-1 antigen gene was very large, or that TIA-1 antigen-related genes were being detected by cross hybridization. A lambda gt11 cDNA derived from RNA-activated T cells was screened with a TIA-1 antigen cDNA probe. Several TIA-1 related cDNAs were identified in this manner. A comparison of these variant clones with TIA-1 genomic sequences indicated that all but one were likely to be splice variants of the TIA-1 gene. Like TIA-1 antigen, TIA1 is a member of a family of RNA-binding proteins, possessing 3 RNA-binding domains and a carboxy terminal auxiliary domain. A segment of the cDNA is useful as a probe for determining the presence of cytolytic lymphocytes in a biological sample. This ability provides an early warning of the presence of an infective agent, e.g. HIV virus.  
See also R32009-17.  
Sequence 375 AA;

**Query Match** 44.4%; Score 60; DB 6; Length 375;  
**Best Local Similarity** 31.6%; Pred. No. 3\_53e+01;  
**Matches** 6; Conservative 9; Mismatches 3; Indels 1; Gaps 1;

**CC**

**Db** 241 gysfvrfnshssahahivs 257  
**ID** W54351; standard; protein; 465 AA.  
**AC** W54351;  
**DT** 14-AUG-1998 (first entry)  
**DE** Vimentin.  
**KW** Endometrium; hyperplasia; adenocarcinoma; proliferative phase; 2D gel electrophoresis; detection.  
**OS** Homo sapiens.  
**PN** WO9810291A1.  
**PD** 12-MAR-1998.  
**PF** 05-SEP-1997; G02394.  
**PR** 08-APR-1997; GB-007132.  
**DR** 06-SEP-1995; GB-018600.  
**PT** (CLIN-) CENT CLINICAL & BASIC RES.  
**PI** Byrjalsen I, Fey SJ, Larsen P;  
**DR** WPI; 98-207057/18.  
**PT** Biochemical markers of human endometrium - useful for, e.g. diagnosis of hyperplasia and adenocarcinoma

**PT** Disclosure: Page 20; 77PP; English.

**CC** Proteins W54349/W54364 are examples of proteins produced in the endometrium during the hyperplasia, adenocarcinoma or proliferative phase of the endometrium. The presence and quantities of these proteins can be detected using 2D gel electrophoresis and comparison of cell lysates. The proteins can be used as biochemical markers to detect the phase of the endometrium and can be measured in body fluids, obviating the need for endometrial biopsies.

**SQ** Sequence 465 AA;

**Query Match** 44.4%; Score 60; DB 31; Length 465;  
**Best Local Similarity** 35.7%; Pred. No. 3\_53e+01;  
**Matches** 5; Conservative 6; Mismatches 3; Indels 0; Gaps 0;

**CC**

**Db** 231 atikkihheeelq 244  
**ID** W24095; standard; Protein; 557 AA.  
**AC** W24095;  
**DT** 20-NOV-1997 (first entry)  
**DE** Enzyme for production of catechol or Pyrogalol compounds.  
**KW** Catalysis; enzyme; methyl group; vanillic acid; syringa; production; catechol; pyrogalol; Pseudomonas paucimobilis.  
**OS** Pseudomonas paucimobilis SYK-6.

**PT** Anderson PJ, Schlossman SF, Streuli M;  
**DR** WPI; 93-045515/05.  
**N-PSDB:** 024556.

**PT** DNA encoding TIA-1 antigen - used for detecting and cytotoxic killing of cytolytic lymphocyte(s) applicable for cancer

**PT** treatment of cytolytic lymphocyte(s) applicable for cancer

**PT** Disclosure: Page 35; 61PP; English.

PN J09173074-A.  
PD 08-JUL-1997.  
PF 21-DEC-1995; 349914.  
PR 21-DEC-1995; JP-349914.

PA (MAZN ) COSMO OIL CO LTD.  
PA (COSM-) COSMO SOGO KENKYUSHO KK.  
DR WPI; 97-397032/37.

DR N-PSDB; T85645.  
PT Gene encoding enzyme which catalyses release of methyl group from  
PT vanillic or syringa acid - useful for production of catechol or  
PT pyrogaryl containing molecules

PS Claim 2; Page 8; 14PP; Japanese.

CC This is an enzyme which is capable of catalysing the reaction of  
CC vanillic acid or syringa acid. The DNA and  
CC transformed host cells can be used to produce a recombinant enzyme. This  
CC enzyme plays an important role in transforming natural components into  
SO those of catechol or pyrogaryl structure.

Sequence 557 AA;

every Match 43.7%; score 59; DB 24; Length 557;  
best Local Similarity 36.8%; Pred. No. 4.48e+01;  
Matches 7; Conservative 6; Mismatches 6; Indels 0; Gaps 0;

Db 539 vpsadfklddeqggqqlf 557  
Qy ||| ||| ::::| |:  
1 VPSYFSFIRTAHDSEVQPLI 19

Search completed: Tue Jan 11 15:41:38 2000  
Job time : 21 secs.



CC LIBRARY: BRAINNOT03  
 CC CLONE: 530522  
 SQ SEQUENCE 344 AA; 38258 MW; 685411 CN;  
 Query Match 45.2%; Score 61; DB 2; Length 344;  
 Best Local Similarity 53.3%; Pred. No. 1; 3.8e+01;  
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 Qy 4 YSFIRTAHDSEVQDL 18

RESULT 2 ID US-08-726-306A-183 STANDARD; PRT; 160 AA.  
 ID US-08-689-974-3 STANDARD; PRT; 494 AA.  
 XX AC XXXXX  
 XX DT DT  
 XX Sequence 3, Application US/08689974  
 DE Sequence 183, Application US/08726306A  
 XX Sequence 183, Application US/08726306A  
 CC Patent No. 5956684  
 CC GENERAL INFORMATION:  
 APPLICANT: van Leeuwen, Frederik Willem  
 APPLICANT: Burbach, Johannes Peter Henri  
 APPLICANT: Grosfeld, Franklin G.  
 TITLE OF INVENTION: DIAGNOSIS METHOD AND REAGENTS  
 NUMBER OF SEQUENCES: 189  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Banner & Witcoff, Ltd.  
 STREET: 1 Financial Center  
 CITY: Boston  
 STATE: MA  
 COUNTRY: US  
 ZIP: 02111  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette, 3.50 inch, 1.44 Mb storage  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: WordPerfect 6.1  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/726,306A  
 FILING DATE: 02-Oct-1996  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: GB 95/20080.4  
 FILING DATE: 02-Oct-1995  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/009,832  
 FILING DATE: 01-Jan-1996  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Williams, Ph.D., Kathleen M.  
 REGISTRATION NUMBER: 34,380  
 REFERENCE/DOCKET NUMBER: 96,048-A (3255/00784)  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 345-9100  
 TELEFAX: (617) 345-9111  
 INFORMATION FOR SEQ ID NO: 183:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 150 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: unknown  
 MOLECULE TYPE: peptide  
 IMMEDIATE SOURCE:  
 LIBRARY: GenBank  
 CLONE: 1293563  
 SQ SEQUENCE 494 AA; 54746 MW; 1410529 CN;  
 Query Match 42.2%; Score 57; DB 2; Length 494;  
 Best Local Similarity 53.3%; Pred. No. 3.61e+01;  
 Matches 8; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 Db 350 YSSIRTSNSQVNL 364  
 Qy 4 YSFIRTAHDSEVQDL 18

RESULT 4 ID US-08-690-734A-96 STANDARD; PRT; 366 AA.  
 XX AC XXXXX  
 XX DT DT  
 XX Sequence 96, Application US/08690734A  
 XX Sequence 96, Application US/08690734A

Query Match 42.2%; Score 57; DB 2; Length 160;  
 Best Local Similarity 35.7%; Pred. No. 3.61e+01;  
 Matches 5; Conservative 5; Mismatches 4; Indels 0; Gaps 0;  
 Db 7 SPLKVKHEEEAEL 20  
 Qy 4 YSFIRTAHDSEVQDL 18

PATENT NO. 5871920  
 GENERAL INFORMATION:  
 APPLICANT: Page, David C.  
 APPLICANT: Reijo, Renee  
 TITLE OF INVENTION: DAZ: A GENE ASSOCIATED WITH AZOOSPERMIA  
 NUMBER OF SEQUENCES: 96  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.  
 STREET: Two Milkton Drive  
 CITY: Lexington  
 STATE: Massachusetts  
 COUNTRY: US  
 ZIP: 02173

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: FastSEQ Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/469,537A  
 FILING DATE: 06-JUN-1995  
 CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/690,734A  
 FILING DATE: 31-JUL-1996  
 CLASSIFICATION: 435

APPLICATION NUMBER: #1.0, version #1.30

CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US 08/310,429  
 FILING DATE: 22-SEP-1994

ATTORNEY/AGENT INFORMATION:  
 NAME: Granahan, Patricia  
 REGISTRATION NUMBER: 32,227  
 REFERENCE/DOCKET NUMBER: WH194-07A

TELECOMMUNICATION INFORMATION:  
 APPLICATION NUMBER: US 08/310,429  
 TELEPHONE: (617) 861-6340  
 TELEFAX: (617) 861-9540  
 LENGTH: 96:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 366 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single

TOPOLOGY: linear  
 MOLECULE TYPE: Peptide  
 SEQUENCE 366 AA; 41236 MW; 802697 CN;

Query Match 41.5%; score 56; DB 2; Length 366;  
 Best Local Similarity 27.8%; Pred. No. 4.58e+01;  
 Matches 5; Conservative 8; Mismatches 5; Indels 0; Gaps 0;

Db 81 GYGFVSVFVNDVWQKIVG 98  
 Qy 3 SYSFIRTAHDSEVQDLIA 20

RESULT 6  
 ID US-07-857-224B-75 STANDARD; PRT; 266 AA.

Query Match 40.7%; Score 55; DB 2; Length 256;  
 Best Local Similarity 47.1%; Pred. No. 5.80e+01;  
 Matches 8; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

Db 90 NFTERNETENPTVKDQIG 106  
 Qy 5 SFIRTAHDSEVQDLIA 20

DT XX  
 DE Sequence 75, Application US/07857224B  
 XX Sequence 75, Application US/07857224B  
 CC Patent No. 5958784  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Benner, Steven A.  
 CC TITLE OF INVENTION: Predicting Folded Structures of Proteins  
 CC NUMBER OF SEQUENCES: 114  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Steven A. Benner  
 CC STREET: Hadlaubstrasse 151  
 CC CITY: Zurich  
 CC STATE: none  
 CC COUNTRY: Switzerland  
 CC ZIP: (note: this is an international post code) CH-8092

COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5 inch diskette, 1.4 Mb storage  
 COMPUTER: Apple Macintosh  
 OPERATING SYSTEM: Macintosh 7.0  
 SOFTWARE: Microsoft Word  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07/857,224B  
 FILING DATE: 03/25/92  
 CLASSIFICATION: 436

PRIOR APPLICATION DATA: none  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (International) 411 632 2830  
 TELEFAX: (International) 411 262 2437  
 TELE: none  
 INFORMATION FOR SEQ ID NO: 75:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 266  
 TYPE: amino acid  
 TOPOLOGY: linear

MOLECULE TYPE: protein  
 DESCRIPTION: protein  
 ORIGINAL SOURCE:  
 ORGANISM: human  
 FEATURE: Protein kinase; Table 8 Column 85  
 PUBLICATION INFORMATION:  
 AUTHORS:  
 AUTHORS: Hanks, S. K.  
 AUTORS: Quinn, A. M.  
 AUTORS: Hunter, T.  
 TITLE: The protein kinase family  
 JOURNAL: Science  
 VOLUME: 241  
 PAGES: 42-52  
 DATE: 1988  
 SEQUENCE: 266 AA; 30314 MW; 399079 CN;  
 SQ

Query Match 40.7%; Score 55; DB 2; Length 266;  
 Best Local Similarity 47.1%; Pred. No. 5.80e+01;  
 Matches 8; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

Db 92 NFIERNETHNPTVKDILG 108  
 Qy :||| :||: | ||: 5 SFIRT-AHDSEVQDLIA 20

RESULT 7 ID US-08-701-191A-19 STANDARD; PRT; 282 AA.  
 XX  
 AC XXXXXX  
 DT  
 DE  
 XX  
 Sequence 19, Application US/08701191A  
 Patent No. 5912428  
 GENERAL INFORMATION:  
 APPLICANT: Moosa Mohammadi, Joseph Schlessinger,  
 APPLICANT: and Steven R. Hubbard  
 TITLE OF INVENTION: CRYSTALS OF THE TYROSINE KINASE DOMAIN  
 NUMBER OF INVENTION: 41  
 NUMBER OF SEQUENCES: 41  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Lyon & Lyon  
 STREET: 633 West Fifth Street  
 CITY: Los Angeles  
 STATE: California  
 COUNTRY: U.S.A.  
 ZIP: 90011-2006  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07-743,357  
 FILING DATE: 21-AUG-1991  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCM/CA90/00062  
 FILING DATE: 23-FEB-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Galle, Edwin J.  
 REGISTRATION NUMBER: 28,584  
 REFERENCE/DOCKET NUMBER: 30924-2

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (613) 237-6900  
 TELEFAX: (613) 237-0045  
 INFORMATION FOR SEQ ID NO: 9:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1003 amino acids  
 TYPE: amino acid

FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Warburg, Richard J.  
 REGISTRATION NUMBER: 32,327  
 REFERENCE/DOCKET NUMBER: 227/088  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (213) 489-1600  
 TELEFAX: (213) 955-0440  
 TELEX: 67-3510  
 INFORMATION FOR SEQ ID NO: 19:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 282 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE: 282 AA; 31925 MW; 445807 CN;  
 SQ

Query Match 40.7%; Score 55; DB 2; Length 282;  
 Best Local Similarity 47.1%; Pred. No. 5.80e+01;  
 Matches 8; Conservative 5; Mismatches 3; Indels 1; Gaps 1;

Db 103 NFIERNETHNPTVKDILG 119  
 Qy :||| :||: | ||: 5 SFIRT-AHDSEVQDLIA 20

RESULT 8 ID US-07-743-357-9 STANDARD; PRT; 1003 AA.  
 XX  
 AC XXXXXX  
 DT  
 DE  
 XX  
 Sequence 9, Application US/07743357  
 Patent No. 5058646  
 GENERAL INFORMATION:  
 APPLICANT: Kang, Yong C.  
 TITLE OF INVENTION: Polypeptide having immunological  
 TITLE OF INVENTION: activity for use as diagnostic reagent and/or vaccine  
 NUMBER OF SEQUENCES: 22  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: KIRBY EADES GALE BAKER  
 STREET: Box 3432, Station D  
 CITY: Ottawa  
 STATE: Ontario  
 COUNTRY: Canada  
 ZIP: K1M 1H8  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/07-743,357  
 FILING DATE: 21-AUG-1991  
 CLASSIFICATION: 424  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: PCM/CA90/00062  
 FILING DATE: 23-FEB-1990  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Galle, Edwin J.  
 REGISTRATION NUMBER: 28,584  
 REFERENCE/DOCKET NUMBER: 30924-2

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (613) 237-6900  
 TELEFAX: (613) 237-0045  
 INFORMATION FOR SEQ ID NO: 9:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1003 amino acids  
 TYPE: amino acid

CC STRANDEDNESS: not relevant  
 CC TOPOLOGY: not relevant  
 CC MOLECULE TYPE: protein  
 CC HYPOTHETICAL: NO  
 CC FRAGMENT TYPE: internal  
 CC ORIGINAL SOURCE:  
 CC ORGANISM: Human immunodeficiency virus type 1  
 CC STRAIN: MAL  
 CC SEQUENCE 1003 AA; 113673 MW; 5160362 CN;  
 SQ Best Local Similarity 40.7%; Score 55; DB 2; Length 1003;  
 Matches 6; Conservative 6; Mismatches 3; Indels 0; Gaps 0;  
 Db 508 YARTKSAHTNDVKQL 522  
 Qy 4 YSFIRTAHDSEVQDL 18

RESULT 9 ID US-08-553-633A-2 STANDARD; PRT; 244 AA.  
 DT DE Sequence 2, Application US/08553633A  
 XX DE Sequence 2, Application US/08553633A  
 XX DE Sequence 2, Application US/08737825  
 XX DE Sequence 2, Application US/08737825  
 CC GENERAL INFORMATION:  
 CC APPLICANT: SALMOND, GEORGE PC  
 CC APPLICANT: HOLDEN, MATTHEW TG  
 CC APPLICANT: COX, ANTHONY RJ  
 CC APPLICANT: THOMSON, NICHOLAS R  
 CC APPLICANT: MCGOWAN, SIMON J  
 CC TITLE OF INVENTION: PROCESS FOR ACTIVATING GENE EXPRESSION  
 CC NUMBER OF INVENTION: IN BACTERIA  
 CC NUMBER OF SEQUENCES: 5  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH  
 CC STREET: PO BOX 747  
 CC CITY: FALLS CHURCH  
 CC STATE: VA  
 CC ZIP: 22040-0747  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: MICROSOFT WORD97  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/553,633A  
 CC FILING DATE: 01-APR-1996  
 CC CLASSIFICATION: 435  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: SVENSSON, LEONARD R.  
 CC SOFTWARE: Patentin Release #1.0, Version #1.30  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/553,633A  
 CC FILING DATE: 01-APR-1996  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (703) 205-8000  
 CC TELEX: (703) 205-8050  
 CC INFORMATION FOR SEQ ID NO: 2:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 244 amino acids  
 CC TYPE: amino acid  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC HYPOTHETICAL: NO  
 CC ANTI-SENSE: NO  
 CC ORIGINAL SOURCE:  
 CC ORGANISM: *Erwinia carotovora*  
 CC SEQUENCE 244 AA; 28137 MW; 305391 CN;  
 SQ Best Local Similarity 40.0%; Score 54; DB 2; Length 244;  
 Matches 6; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 Db 98 SGYTFVLHDHNNVATL 114  
 Qy 2 PSYSFIRTAHDSEVQDL 18

RESULT 10 ID US-08-737-825-2 STANDARD; PRT; 244 AA.  
 DT DE Sequence 2, Application US/08737825  
 XX DE Sequence 2, Application US/08737825  
 XX DE Sequence 2, Application US/08737825  
 CC GENERAL INFORMATION:  
 CC APPLICANT: SALMOND, GEORGE PEACOCK COPELAND  
 CC APPLICANT: MCGOWAN, SIMON JAMES  
 CC APPLICANT: SEBAHIA, MOHAMMED  
 CC APPLICANT: COX, ANTHONY RICHARD JOHN  
 CC APPLICANT: HOLDEN, MATTHEW THOMAS GEOFFREY  
 CC APPLICANT: BORDER, LAUREN ELIZABETH  
 CC APPLICANT: BYCROFT, BARRIE WALSHAM  
 CC APPLICANT: WILLIAMS, PAUL  
 CC TITLE OF INVENTION: GENES INVOLVED IN THE BIOSYNTHETIC PATHWAY OF CARBAPENEM  
 CC NUMBER OF SEQUENCES: 12  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Birch, Stewart, Kolasch & Birch, LLP  
 CC STREET: P.O. Box 747  
 CC CITY: Falls Church  
 CC STATE: Virginia  
 CC COUNTRY: USA  
 CC ZIP: 22040-0747  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: MICROSOFT WORD97  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/737,825  
 CC FILING DATE: 03-JUN-1997  
 CC CLASSIFICATION: 435  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: SVENSSON, LEONARD R.  
 CC REGISTRATION NUMBER: 30,330  
 CC REFERENCE/DOCKET NUMBER: 1009-0105P  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (703) 205-8000  
 CC TELEX: (703) 205-8050  
 CC INFORMATION FOR SEQ ID NO: 2:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 244 amino acids  
 CC TYPE: amino acid  
 CC STRANDEDNESS:  
 CC TOPOLOGY: linear  
 CC MOLECULE TYPE: protein  
 CC HYPOTHETICAL: NO  
 CC ANTI-SENSE: NO  
 CC ORIGINAL SOURCE:  
 CC ORGANISM: *Erwinia carotovora*  
 CC SEQUENCE 244 AA; 28137 MW; 305391 CN;  
 SQ Best Local Similarity 35.3%; Score 54; DB 2; Length 244;  
 Matches 6; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 Db 98 SGYTFVLHDHNNVATL 114  
 Qy 2 PSYSFIRTAHDSEVQDL 18

Query Match 40.0%; Score 54; DB 2; Length 244;  
 CC Best Local Similarity 35.3%; Pred. No. 7.33e+01;  
 Matches 6; Conservative 6; Mismatches 5; Indels 0; Gaps 0;  
 Sq 98 SGYTFVLHDHNNVATL 114

Query Match 40.0%; Score 54; DB 2; Length 244;  
 CC Best Local Similarity 35.3%; Pred. No. 7.33e+01;  
 Matches 6; Conservative 6; Mismatches 5; Indels 0; Gaps 0;

**RESULT** 11  
**ID** US-08-746-682A-13    **STANDARD;**    **PRT;** 473 AA.  
**XX**  
**AC** XXXXXX  
**XX**  
**DT**  
**DE**  
**XX**  
**Sequence 13, Application US/08746682A**  
**Patent No. 5786584**  
**GENERAL INFORMATION:**  
**APPLICANT:** STINGELE, Francesca  
**APPLICANT:** MOLLET, Beat  
**TITLE OF INVENTION:** LACTIC BACTERIA PRODUCING  
**TITLE OF INVENTION:** EXOPOLYSACCHARIDES  
**NUMBER OF SEQUENCES:** 19  
**CORRESPONDENCE ADDRESS:**  
**ADDRESSEE:** Pennie & Edmonds  
**STREET:** 1155 Avenue of the Americans  
**CITY:** New York  
**STATE:** New York  
**ZIP:** 10036  
**COMPUTER READABLE FORM:**  
**MEDIUM TYPE:** FLOPPY disk  
**COMPUTER:** IBM PC compatible  
**OPERATING SYSTEM:** PC-DOS/MS-DOS  
**SOFTWARE:** Patent Release #1.0, Version #1.30  
**CURRENT APPLICATION DATA:**  
**ADDRESSEE:** Pennie & Edmonds  
**STREET:** 1155 Avenue of the Americans  
**CITY:** New York  
**STATE:** New York  
**ZIP:** U.S.A.  
**CLASSIFICATION:** 435  
**PRIOR APPLICATION DATA:**  
**APPLICATION NUMBER:** US/08/746,682A  
**FILING DATE:** 14-NOV-1995  
**CURRENT APPLICATION DATA:**  
**APPLICATION NUMBER:** 08/597,236  
**FILING DATE:** 20-JUN-1995  
**APPLICATION NUMBER:** EP 95201669.9  
**FILING DATE:** 20-JUN-1995  
**ATTORNEY/AGENT INFORMATION:**  
**NAME:** Fanucci A., Allan  
**REGISTRATION NUMBER:** 30256  
**INFORMATION FOR SEQ ID NO: 13:**  
**SEQUENCE CHARACTERISTICS:**  
**MOLECULE TYPE:** protein  
**LENGTH:** 473 amino acids  
**TYPE:** amino acid  
**TOPOGY:** linear  
**SEQUENCE:** 473 AA; 53028 MW; 1298864 CN;  
**SEQUENCE CHARACTERISTICS:**  
**MOLECULE TYPE:** protein  
**LENGTH:** 473 amino acids  
**TYPE:** amino acid  
**TOPOGY:** linear  
**SEQUENCE:** 473 AA; 53028 MW; 1298864 CN;  
**Query Match** 40.0%; **Score** 54; **DB** 1; **Length** 473;  
**Best Local Similarity** 33.3%; **Pred** No. 7.33e+01;  
**Matches** 6; **Conservative** 6; **Mismatches** 6; **Indels** 0; **Gaps** 0;  
**Db** 29 VPLVTTAMTPQEYGMADL 46  
**Oy** 1 VPSYFIRTAHSEVQDL 18  
**RESULT** 13  
**ID** US-08-162-081B-37    **STANDARD;**    **PRT;** 1069 AA.  
**XX**  
**AC** XXXXXX  
**XX**  
**DT**  
**DE**  
**XX**  
**Sequence 37, Application US/08162081B**  
**Sequence 37, Application US/08162081B**  
**Patent No. 5824492**  
**GENERAL INFORMATION:**  
**APPLICANT:** Hiles, Ian Donald; Fry, Michael John; Dhund, Ritt  
**APPLICANT:** Bala, Waterfield, Michael Derek; Parker, Peter  
**APPLICANT:** Jospah, Otsu; Masayuki; Panayotou, George; Wolin  
**APPLICANT:** Stepano, Gout, Ivan Tarassovitch  
**TITLE OF INVENTION:** POLYPEPTIDES HAVING KINASE ACTIVITY,  
**TITLE OF INVENTION:** THEIR PREPARATION AND USE  
**NUMBER OF SEQUENCES:** 50  
**CORRESPONDENCE ADDRESS:**  
**ADDRESSEE:** Feifei & Lynch  
**STREET:** 805 Third Avenue  
**CITY:** New York  
**STATE:** New York  
**ZIP:** 10022  
**COMPUTER READABLE FORM:**  
**MEDIUM TYPE:** Diskette, 5.25 inch, 360 kb storage

CC COMPUTER: IBM PS/2  
 CC OPERATING SYSTEM: PC-DOS  
 CC SOFTWARE: Wordperfect  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/162,081B  
 CC FILING DATE: February 7, 1994  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: PCT/GB93/00761  
 CC FILING DATE: 13 April 1993  
 CC ATTORNEY/AGENT INFORMATION:  
 NAME: Pasqualini, Patricia A.  
 REGISTRATION NUMBER: 34,894  
 REFERENCE/DOCKET NUMBER: LUD 5256  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 688-9200  
 TELEFAX: (212) 838-3884  
 INFORMATION FOR SEQ ID NO: 37:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1069 amino acids  
 LENGTH: 1069 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLogy: linear  
 STRANDEDNESS: single  
 TOPOLogy: linear  
 SQ SEQUENCE 1069 AA; 124437 MW; 5707845 CN;  
 Query Match 40.0%; Score 54; DB 2; Length 1069;  
 Best Local Similarity 22.2%; Pred. No. 7.33e+01;  
 Matches 4; Conservative 10; Mismatches 4; Indels 0; Gaps 0;  
 Db 66 SSYIVSVTQEAERBEFF 83  
 Qy :|||: ::::| :::::  
 2 PSYSFIRTAHDSEVQDLI 19  
 RESULT 14 ID US-08-780-872-37 STANDARD; PRT; 1069 AA.  
 XX AC XXXXXX  
 CC GENERAL INFORMATION:  
 DT APPLICANT: Hiles, Ian Donald; Fry, Michael John; Dhand, Ritu  
 CC APPLICANT: Bala, Waterfield, Michael Derek; Parker, Peter  
 CC APPLICANT: Joseph, Otsu, Masayuki; Papavayou, George; Volinia,  
 DE Sequence 37, Application US/08780872  
 XX Sequence 37, Application US/08780872  
 CC Patent No. 5846824  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Hiles, Ian Donald; Fry, Michael John; Dhand, Ritu  
 CC APPLICANT: Bala, Waterfield, Michael Derek; Parker, Peter  
 CC APPLICANT: Joseph, Otsu, Masayuki; Papavayou, George; Volinia,  
 CC APPLICANT: Stefano, Gout, Ivan Tarasovitch  
 CC TITLE OF INVENTION: POLYPEPTIDES HAVING KINASE ACTIVITY,  
 CC NUMBER OF SEQUENCES: 50  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Felte & Lynch  
 CC STREET: 805 Third Avenue  
 CC CITY: New York  
 CC STATE: New York  
 CC COUNTRY: USA  
 CC ZIP: 10022  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
 CC COMPUTER: IBM PS/2  
 CC OPERATING SYSTEM: PC-DOS  
 CC SOFTWARE: Wordperfect  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/780,872  
 CC FILING DATE: 0-JAN-1997  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: 08/162,081  
 CC FILING DATE: February 7, 1994  
 CC ATTORNEY/AGENT INFORMATION:  
 NAME: Pasqualini, Patricia A.  
 REGISTRATION NUMBER: 34,894  
 REFERENCE/DOCKET NUMBER: LUD 5256  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 688-9200  
 TELEFAX: (212) 838-3884  
 INFORMATION FOR SEQ ID NO: 37:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1069 amino acids  
 LENGTH: 1069 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLogy: linear  
 STRANDEDNESS: single  
 TOPOLogy: linear  
 SQ SEQUENCE 1069 AA; 124437 MW; 5707845 CN;  
 Query Match 40.0%; Score 54; DB 2; Length 1069;  
 Best Local Similarity 22.2%; Pred. No. 7.33e+01;  
 Matches 4; Conservative 10; Mismatches 4; Indels 0; Gaps 0;  
 Db 66 SSYIVSVTQEAERBEFF 83  
 Qy :|||: ::::| :::::  
 2 PSYSFIRTAHDSEVQDLI 19  
 RESULT 15 ID US-08-780-872-36 STANDARD; PRT; 1080 AA.  
 XX AC XXXXXX  
 CC Sequence 36, Application US/08780872  
 CC Patent No. 5846824  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Hiles, Ian Donald; Fry, Michael John; Dhand, Ritu  
 CC APPLICANT: Bala, Waterfield, Michael Derek; Parker, Peter  
 CC APPLICANT: Joseph, Otsu, Masayuki; Papavayou, George; Volinia,  
 CC APPLICANT: Stefano, Gout, Ivan Tarasovitch  
 CC TITLE OF INVENTION: POLYPEPTIDES HAVING KINASE ACTIVITY,  
 CC NUMBER OF SEQUENCES: 50  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Felte & Lynch  
 CC STREET: 805 Third Avenue  
 CC CITY: New York  
 CC STATE: New York  
 CC COUNTRY: USA  
 CC ZIP: 10022  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Diskette, 5.25 inch, 360 kb storage  
 CC COMPUTER: IBM PS/2  
 CC OPERATING SYSTEM: PC-DOS  
 CC SOFTWARE: Wordperfect  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/780,872  
 CC FILING DATE: 0-JAN-1997  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: 08/162,081  
 CC FILING DATE: February 7, 1994  
 CC ATTORNEY/AGENT INFORMATION:  
 NAME: Pasqualini, Patricia A.  
 REGISTRATION NUMBER: 34,894  
 REFERENCE/DOCKET NUMBER: LUD 5256  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (212) 688-9200  
 TELEFAX: (212) 838-3884  
 INFORMATION FOR SEQ ID NO: 36:

SEQUENCE CHARACTERISTICS:  
CC LENGTH: 1080 amino acids  
CC TYPE: amino acid  
CC STRANDEDNESS: single  
CC TOPOLOGY: linear  
SQ SEQUENCE 1080 AA; 125733 MW; 5831251 CN;  
Query Match 40.0%; Score 54; DB 2; Length 1080;  
Best Local Similarity 22.2%; Pred. No. 7.33e+01;  
Matches 4; Conservative 10; Mismatches 4; Indels 0; Gaps 0;  
Db 66 SSYIFVSVTQEEREFF 83  
Qy 1 :||| :|| :||| :||| :|||  
2 PSYSFIRTAHDSEVQDLI 19

Search completed: Tue Jan 11 15:46:11 2000  
time : 8 secs.



ALTERNATE\_NAMES sucrose 6-glucosyltransferase  
 ORGANISM #formal\_name Streptococcus mutans  
 DATE 31-Mar-1992 #sequence\_revision 31-Mar-1992 #text\_change  
 ACCESSIONS JTM0345; C33135  
 REFERENCES Ueda, S.; Shiroza, T.; Kuramitsu, H.K.  
 #authors 1  
 #Journal Genie (1988) 69:101-109  
 #cross-references MUID:89137980  
 #accession JTM0345  
 #title #molecule\_type DNA  
 #residues 1-1375 #label UED  
 #experimental\_source GS-5  
 REFERENCE A33135  
 #authors Shiroza, T.; Ueda, S.; Kuramitsu, H.K.  
 #Journal J. Bacteriol. (1987) 169:4263-4270  
 #title Sequence analysis of the gtfB gene from *Streptococcus mutans*.  
 #cross-references MUID:87308013  
 #accession C33135  
 #status preliminary  
 #molecule\_type DNA  
 #residues 1-349 #label SHI  
 GENETICS #cross-references GB:MI17361  
 #gene gtfC  
 FUNCTION catalyzes the synthesis of both water-soluble and  
 #description water-insoluble glucans from glucose.  
 CLASSIFICATION #superfamily cpl repeat homology/  
 #keywords duplication; glycosyltransferase; hexosyltransferase;  
 FEATURE transferase  
 1-34 #domain signal sequence #status predicted #label SIG\\  
 35-1375 #product glucosyltransferase #status predicted #label  
 MAT\\  
 1126-1145 #domain cpl repeat homology #label CPL\\  
 1253-1272 #domain cpl repeat homology #label CPL2\\  
 1318-1337 #domain cpl repeat homology #label CPL3  
 SUMMARY #length 1375 #molecular-weight 153021 #checksum 7015  
 Query Match 82.2%; Score 111; DB 2; Length 1375;  
 Best Local Similarity 94.7%; Pred. No. 3.42e-10; Matches 18; Conservative 0; Mismatches 0; Indels 1; Gaps 1;  
 Db 578 VPSYSFIR-AHDSEVQDLI 595  
 1 VPSYSFIRAHDSEVQDLI 19  
 RESULT 3  
 ENTRY A38175 #type complete  
 TITLE glucosyltransferase precursor - *Streptococcus sobrinus*  
 ORGANISM #formal\_name *Streptococcus sobrinus*  
 DATE 28-Aug-1992 #sequence\_revision 28-Aug-1992 #text\_change  
 ACCESSIONS A38175  
 REFERENCES Abo, H.; Matsuura, T.; Kodama, T.; Ohta, H.; Fukui, K.; Kato, J.; Bacteriol. (1991) 173:999-995  
 #authors 1  
 #Journal Peptide sequences for sucrose splitting and glucan binding  
 #title within *Streptococcus sobrinus* glucosyltransferase  
 (#water-insoluble glucan synthetase).  
 #cross-references MUID:91123227  
 #accession A38175  
 #status preliminary  
 #molecule\_type DNA  
 #residues 1-1592 #label ABO  
 #cross-references GB:D90213; NTID:g9217032; PID:d1014946; PID:g9217033  
 #classification superfamily cpl repeat homology  
 FEATURE #cross-references MUID:90316665

SUMMARY #length 1592 #molecular\_weight 176167 #checksum 5940  
 Query Match 71.9%; Score 97; DB 2; Length 1592;  
 Best Local Similarity 84.2%; Pred. No. 4.36e-07; Matches 16; Conservative 1; Mismatches 1; Indels 1; Gaps 1;  
 Db 548 VPSYFAR-AHDSEVQDLI 565  
 QY 1 VPSYSFIRAHDSEVQDLI 19  
 RESULT 4  
 ENTRY JC5473 #type complete  
 TITLE dextranucrase (EC 2.4.1.5) - *Leuconostoc mesenteroides*  
 ORGANISM #formal\_name *Leuconostoc mesenteroides*  
 DATE 07-Jul-1997 #sequence\_revision 29-Aug-1997 #text\_change  
 ACCESSIONS JC5473  
 REFERENCES JC5473  
 #authors Monsan, P.  
 #journal Gene (1996) 182:23-32  
 #title Cloning and sequencing of a gene coding for a novel  
 dextranucrase from *Leuconostoc mesenteroides* NRRL B-1299  
 #cross-references MUID:97136686  
 #accession JC5473  
 #status nucleic acid sequence not shown; translation not shown  
 #molecule\_type DNA  
 #residues 1-1290 #label MON  
 #cross-references GB:U8181  
 COMMENT This enzyme catalyses the transfer of D-glucopyranosyl units from  
 sucrose onto acceptor molecules.  
 GENETICS #gene dsRA  
 #keywords glycosyltransferase; hexosyltransferase  
 FEATURE #domain catalytic #status predicted #label CAT\\  
 78-870 #domain glucan-binding #status predicted #label GCB  
 SUMMARY #length 1290 #molecular\_weight 145572 #checksum 9386  
 Query Match 68.1%; Score 92; DB 2; Length 1290;  
 Best Local Similarity 78.9%; Pred. No. 5.15e-06; Matches 15; Conservative 2; Mismatches 1; Indels 1; Gaps 1;  
 Db 388 PNYSFIR-AHDSEVTTIA 405  
 QY 2 PSYSFIRAHDSEVODLIA 20  
 RESULT 5  
 ENTRY A41483 #type complete  
 TITLE glucosyltransferase (EC 2.4.1.-) qtfS precursor -  
 ORGANISM *Streptococcus sobrinus*  
 DATE #formal\_name *Streptococcus sobrinus*  
 ACCESSIONS 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change  
 REFERENCES A41483  
 #authors Gilmore, R.S.; Russell, R.R.B.; Ferretti, J.J.  
 #Journal Infect. Immun. (1990) 58:2452-2458  
 #title Analysis of the *Streptococcus* downy gtfS gene, which  
 specifies a glucosyltransferase that synthesizes soluble  
 glucans.



Query Match 54.8%; Score 74; DB 2; Length 644;  
 Best Local Similarity 56.3%; Pred. No. 5 58e-02;  
 Matches 9; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 Db 17 AFLRNHHEEVADLLA 32  
 Qy 5 SFIRTAHDSEVQDLIA 20

RESULT 10  
 ENTRY S22737 #type complete  
 TITLE glucosyltransferase (EC 2.4.1.-) S - Streptococcus salivarius  
 ORGANISM #formal\_name Streptococcus salivarius  
 DATE 16-Sep-1992 #sequence\_revision 16-Sep-1992 #text\_change  
 09-Sep-1997  
 ACCESSIONS S22737; S28810; B44811; S22727  
 REFLIST RENE  
 #authors J. Gen. Microbiol. (1991) 137:557-593  
 #submission submitted to the EMBL Data Library, March 1992  
 #accession S22737  
 #molecule\_type DNA  
 #residues 1-1599 #label JNC  
 #cross-references EMBL:21182; NID:947530; PID:947531  
 REFERENCE A44811  
 #authors Giffard, P.M.; Simpson, C.L.; Milward, C.P.; Jacques, N.A.  
 #journal Molecular characterization of a cluster of at least two  
 glucosyltransferase genes in Streptococcus salivarius ATCC  
 25975  
 #cross-references NID:92148377  
 #accession S28810  
 #molecule\_type DNA  
 #residues 1-51 #label GIF  
 #cross-references EMBL:211873  
 GENETICS  
 #gene gtfk  
 CLASSIFICATION #superfamily cpl repeat homology  
 #keywords glycosyltransferase; hexosyltransferase  
 SUMMARY #length 1599 #molecular\_weight 176480 #checksum 3135  
 Query Match 54.8%; Score 74; DB 2; Length 1599;  
 Best Local Similarity 63.2%; Pred. No. 2.32e-02;  
 Matches 12; Conservative 4; Mismatches 2; Indels 1; Gaps 1;  
 ACCESSIONS 573 ATLFVVR-HDSEVQTVIA 590  
 Qy 2 PSVSFIRTAHDSEVQDLIA 20

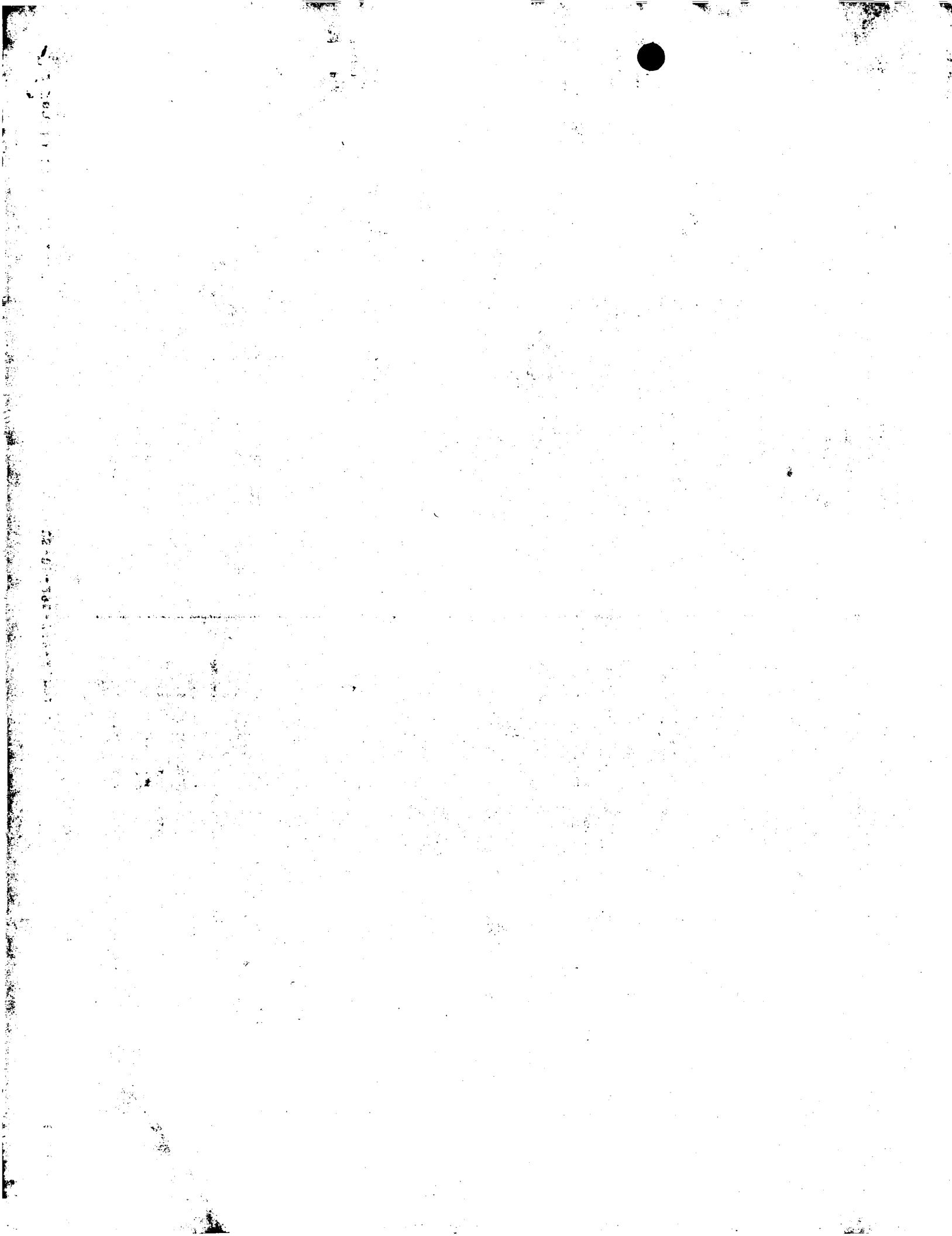
RESULT 11  
 ENTRY S51349 #type complete  
 TITLE hypothetical protein YLR346c - yeast (saccharomyces  
 cerevisiae)  
 ALTERNATE\_NAMES hypothetical protein I8300\_14  
 ORGANISM #formal\_name Saccharomyces cerevisiae  
 DATE 23-Feb-1995 #sequence\_revision 12-May-1995 #text\_change  
 21-Nov-1997  
 ACCESSIONS S51349  
 REFERENCE S51338  
 #authors Du, Z.  
 #submission submitted to the EMBL Data Library, December 1994  
 #description The sequence of S. cerevisiae cosmid 8300.  
 #accession S51349  
 #molecule\_type DNA  
 #residues 1-101 #label DIZ  
 #cross-references EMBL:U19028; NID:9609380; PID:9609393; MIPS:YLR346C  
 #map\_position 12R  
 SUMMARY #length 101 #molecular\_weight 11348 #checksum 3891

RESULT 12  
 ENTRY S05075 #type fragments  
 TITLE neurofilament triplet M protein - pig (fragments)  
 ALTERNATE NAMES 16k neurofilament protein; NF-M(medium) protein; type IV IFP  
 PROTEIN protein  
 ORGANISM #formal\_name Sus scrofa domestica #common\_name domestic pig  
 DATE 31-Mar-1988 #sequence\_revision 02-Jul-1998 #text\_change  
 10-Jul-1998  
 ACCESSIONS A05075; S02510  
 REFERENCE Geisler, N.; Fischer, S.; Vandekerckhove, J.; Plessmann, U.;  
 Weber, K.  
 #authors EMBL J. (1984) 3:2701-2706  
 #journal Hybrid character of a large neurofilament protein (NF-M):  
 Intermediate filament type sequence followed by a long and  
 acidic carboxy-terminal extension.  
 #cross-references NID:85076594  
 #accession A05075  
 #molecule\_type protein  
 #residues 1-454 #label GEI  
 #experimental\_source spinal cord  
 REFERENCE S02570  
 #authors Gaisler, N.; Vandekerckhove, J.; Weber, K.  
 #journal FEBS Lett. (1987) 221:403-07  
 #title Location and sequence characterization of the major  
 phosphorylation sites of the high molecular mass  
 neurofilament proteins M and H.  
 #cross-references NID:87304852  
 #accession S02570  
 #molecule\_type protein  
 #residues 438-450;455-459;460-475;476-514;515-532 #label GEI2  
 #experimental\_source spinal cord  
 #superfamily cytoskeletal keratin  
 CLASSIFICATION #length 532  
 KEYWORDS blocked amino end; coiled coil; intermediate filament;  
 FEATURE phosphoprotein  
 #domain head #status predicted #label HED  
 1-98 99-412 438-454, 455-459,  
 460-475, 476-514 #domain alpha-helical rod #label ROD  
 438-454, 455-459,  
 460-475, 476-514 #domain tail (fragments) #status predicted #label TIP1  
 515-532 #modified\_site blocked amino end (Ser) (probably  
 acetylated) #status experimental  
 456, 462, 465, 479 #binding\_site phosphate (Ser) (covalent) #status  
 experimental  
 SUMMARY #length 532 #checksum 5520  
 Query Match 52.6%; Score 71; DB 1; Length 532;  
 Best Local Similarity 50.0%; Pred. No. 8.62e-02;  
 Matches 8; Conservative 5; Mismatches 3; Indels 0; Gaps 0;  
 Db 232 AFLRNHHEEVADLLA 247  
 Qy 5 SFIRTAHDSEVQDLIA 20

RESULT 13  
 ENTRY A45669 #type complete  
 TITLE neurofilament triplet M protein - rat  
 ORGANISM #formal\_name Rattus norvegicus #common\_name Norway rat  
 DATE 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change  
 17-Mar-1999  
 ACCESSIONS A45669; S25712; A42393

```

REFERENCE A45669 #cross-references MUID:8815120
#authors Napolitano, E.W.; Chin, S.S.M.; Colman, D.R.; Liem, R.K.H.
#Journal J. Neurosci. (1987) 7:250-259
#title Complete amino acid sequence and in vitro expression of rat
#cross-references MUID:87282018
#accession A45669 #status preliminary; not compared with conceptual translation
#molecule_type mRNA #bindingsite phosphate (Ser) (covalent) #status
##residues 1-845 ##label NAP 615, 628, 641, 654, #predicted
#cross-references EMBL:M18628; NID:g205687; PID:g205688 #molecular_weight 102447 #checksum 8997
REFERENCE S25712 #authors Kelly, B.M.; Gillespie, C.S.; Sherman, D.L.; Brophy, P.J.
#Journal J. Cell Biol. (1992) 118:397-410
#title Schwann cells of the myelin-forming phenotype express
#cross-references MUID:92332596 neurofilament protein NF-M.
#accession S25712 #status preliminary
##molecule_type mRNA #residues 1-17, 19-21, 'P', 23-204, 'L', 206-500, 'E', 501-845 ##label KEL
#cross-references EMBL:212152; NID:g56751; PID:g56752
REFERENCE A42303 #authors Xu, Z.S.; Liu, W.S.; Willard, M.B.
#Journal J. Biol. Chem. (1992) 267:4467-4471
#title Identification of six phosphorylation sites in the
#cross-references M. COOH-terminal tail region of the rat neurofilament protein
#accession A42303 #status preliminary; not compared with conceptual translation
##molecule_type nucleic acid #residues 411-500, 'E', 501-843 ##label XU1
#note sequence extracted from NCBI backbone (NCBIPR:83873)
CLASSIFICATION #superfamily cytoskeletal keratin
#length 845 #molecular_weight 95661 #checksum 1701
SUMMARY Query Match 52.6%; Score 71; DB 2; Length 916;
Best Local Similarity 50.0%; Pred. No. 8.62e-02;
Matches 8; Conservative 5; Mismatches 2; Indels 0; Gaps 0;
Db 232 AF1RSNHEEEVADLLA 247
OY ::::: ::::: ::::: :::::
5 SFIRTAHDSEVQDLIA 20
RESULT 15 #cross-references MUID:92165197
ENTRY 152658 #type complete
TITLE neurofilament-66 - human
ORGANISM #formal_name Homo sapiens #common_name man
DATE 02-Jul-1996 #sequence_revision 02-Jul-1996 #text_change
16-Feb-1997
ACCESSIONS 152658
REFERENCE 152658 #status preliminary; translated from GB/EMBL/DDBJ
#authors Chan, S.O.; Chiu, F.C.
#Journal Brain Res. Mol. Brain Res. (1995) 29:177-184
#title Cloning and developmental expression of human 66 kd
neurofilament protein.
#cross-references MUID:95287809
#accession I52658 #status preliminary; translated from GB/EMBL/DDBJ
##molecule_type mRNA #residues 1-494 ##label RES
##cross-references GB:S73296; NID:g994843; PID:g994844
CLASSIFICATION #superfamily cytoskeletal keratin
SUMMARY #length 494 #molecular_weight 54907 #checksum 4489
Query Match 51.9%; Score 70; DB 2; Length 494;
Best Local Similarity 50.0%; Pred. No. 1.33e-01;
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;
Db 221 AFVRQVIDEEVALLA 236
OY ::::: ::::: ::::: :::::
5 SFIRTAHDSEVQDLIA 20
RESULT 14 #cross-references MUID:87275053
ENTRY A27864 #type complete
TITLE neurofilament triplet M protein - human
ORGANISM NF-M (medium) protein
#formal_name Homo sapiens #common_name man
DATE 30-Jun-1989 #sequence_revision 30-Jun-1989 #text_change
20-Mar-1998
ACCESSIONS A27864; A30157
REFERENCE #authors Myers, M.W.; Lazzarini, R.A.; Lee, V.M.Y.; Schlaepfer, W.W.;
#Journal Nelson, D.L.
#title The human mid-size neurofilament subunit: a repeated protein
sequence and the relationship of its gene to the
intermediate filament gene family.
#cross-references MUID:87275053
#accession A27864 #molecule_type DNA
##residues 1-916 ##label MYE
#cross-references GB:Y00067; NID:g35045; PID:g35046
REFERENCE A30157 #authors Lee, V.M.Y.; Ottos Jr., L.; Carden, M.J.; Hollosi, M.#
#cross-references A27864 Dietzschold, B.; Lazzarini, R.A.
#journal Proc. Natl. Acad. Sci. U.S.A. (1988) 85:1998-2002
#title Identification of the major multiphosphorylation site in
mammalian neurofilaments.
```









FT REPEAT 1225 1274 3.  
 FT REPEAT 1289 1339 4. (INCOMPLETE).  
 FT REPEAT 1353 1365 5.  
 SEQUENCE 1365 AA; 151590 MW; 15688B31 CRC32;

Query Match 63.7%; Score 86; DB 1; Length 1365;  
 Best Local Similarity 75.0%; Pred. No. 1.0e-05;  
 Matches 15; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

QY 1 VPSYSFIRTAHSEVQDIA 20

RESULT 6 ID NFM\_RABBIT STANDARD; PRT; 1430 AA.  
 ID GTFD\_STRMU STANDARD; PRT; 1430 AA.  
 AC PA9331; DT 01-FEB-1996 (REL. 33, CREATED)  
 DT 01-FEB-1996 (REL. 33, LAST ANNOTATION UPDATE)  
 DE GLUCOSYLTRANSFERASE-S PRECURSOR (EC 2.4.1.5) (GTF-S) (DEXTRANSUCRASE)  
 DE (SUCROSE 6-GLUCOSYLTRANSFERASE).  
 GN GTFD.  
 OS SIREPTOCOCCUS MUTANS.  
 OC BACTERIA; FIRMICUTES; BACILLUS/CLOSTRIDIUM GROUP; STREPTOCOCCACEAE;  
 OC SIREPTOCOCCUS.  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAINS-G-5;  
 RX MEDLINE; 91100958.  
 RA HONDA O., KATO C., KURAMITSU H.K.;  
 RA RT  
 RL J. GEN. MICROBIOL. 136: 2099-2110(1990).  
 CC -!- FUNCTION: PRODUCTION OF EXTRACELLULAR GLUCANS, THAT ARE THOUGHT  
 TO PLAY A KEY ROLE IN THE DEVELOPMENT OF THE DENTAL PLAQUE BECAUSE  
 OF THEIR ABILITY TO ADHERE TO SMOOTH SURFACES AND MEDIATE THE  
 AGGREGATION OF BACTERIAL CELLS AND FOOD DEBRIS.  
 CC -!- CATALYTIC ACTIVITY: SUCROSE + (1,6-ALPHA-D-GLUCOSYL)(N) -  
 CC D-FRUCTOSE + (1,6-ALPHA-D-GLUCOSYL)(N+1).  
 CC -!- SUBCELLULAR LOCATION: SECRETED.  
 CC -!- DISEASE: DENTAL CARIES.  
 CC -!- GTF-I SYNTHESIZES WATER-INSOLUBLE GLUCANS (ALPHA 1,3-LINKED  
 CC GLUCOSIDE AND SOME 1,6 LINKAGES); GTF-SI SYNTHESIZES WATER-SOLUBLE  
 CC GLUCANS (ALPHA 1,6-GLUCOSE).  
 CC -!- SIMILARITY: TO OTHER GLUCOSYLTRANSFERASES AND SOME TO A GLUCAN-  
 CC BINDING PROTEIN FROM S. MUTANS.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to license@isb-sib.ch).

DR EMBL; 247378; GS54353; -.  
 DR PFAM; PS00226; IF; 1.  
 DR PFAM; PF00038; filament; 1.  
 KW INTERMEDIATE FILAMENT; HETRAD REPEAT PATTERN; COILED COIL; NEURONE;  
 KW PHOSPHORYLATION; GLYCOPROTEIN.

FT NON\_TER 1  
 FT DOMAIN <1 197 ROD.  
 FT DOMAIN 198 643 TAIL.  
 FT DOMAIN <1 33 COIL 1B.  
 FT DOMAIN 34 50 LINKER 12.  
 FT DOMAIN 51 72 COIL 2A.  
 FT DOMAIN 73 76 LINKER 2.  
 FT DOMAIN 77 197 COIL 2B.  
 FT CARBOHYD 217 GLCNAC (BY SIMILARITY).  
 SQ SEQUENCE 644 AA; 72450 MW; 09A05E7 CRC32;

Query Match 58.5%; Score 79; DB 1; Length 1430;  
 Best Local Similarity 65.0%; Pred. No. 4.0e-04;  
 Matches 13; Conservative 4; Mismatches 2; Indels 1; Gaps 1;

QY 5 SFRTAHSEVQDIA 20

RESULT 8 ID NFM\_PIG STANDARD; PRT; 454 AA.  
 ID AC P08552; DT 01-AUG-1988 (REL. 08, CREATED)  
 DT 01-AUG-1988 (REL. 08, LAST SEQUENCE UPDATE)  
 DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE NEUROFILAMENT TRIPLET M PROTEIN (160 KD NEUROFILAMENT PROTEIN) (NF-M)  
 DE (FRAGMENT).  
 DE  
 DE NFM.  
 DE SUS SCROFA (PIG).  
 DE EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 DE ARTIODACTyla; SUIFORMES; SUINA; SUIDAE; SUS.  
 DE [1]  
 DE RN  
 DE RP  
 DE SEQUENCE.  
 DE TISSUE-SPECIFIC.  
 DE  
 DE MEDLINE; 85076594.  
 DE  
 DE RA GEISLER N., FISCHER S., VANDEKERCKHOVE J., PLESSMANN U., WEBER K.;  
 DE RT EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;  
 DE RT intermediate filament type sequence followed by a long and acidic  
 DE RT carboxy-terminal extension.;"  
 DE RL EMBO J. 3:2701-2706(1984).  
 DE -!- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,  
 DE AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.  
 DE -!- PRT: THERE ARE A NUMBER OF REPEATS OF THE TRIPETIDE K-S-P, NFN M I  
 DE PROPHOSPHORYLATED ON A NUMBER OF THE SERINES IN THIS MOTIF. IT IS  
 DE THOUGHT THAT PHOSPHORYLATION OF NFH RESULTS IN THE FORMATION OF  
 DE INTERFILAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTENANCE  
 DE OF AXONAL CALIBER.  
 DE CC -!- PTM: PHOSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTIONING  
 DE OF THE LARGER NEUROFILAMENT POLYPEPTIDES (NF-M AND NF-H), THE  
 DE LEVELS OF PHOSPHORYLATION BEING ALTERED DEVELOPMENTALLY AND  
 DE COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION.  
 DE CC -!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.  
 DE DR PIR; A05075; QFCM.  
 DE PROSITE: PS00226; IF; 1.  
 DE DR PF00038: filament; 1.  
 DE KW INTERMEDIATE FILAMENT; HEPPAD REPEAT PATTERN; COILED COIL; NEURONE;  
 DE ACETYLATION; PHOSPHORYLATION; GLYCOPROTEIN.  
 DE MOD RSS 1  
 DE FT DOMAIN 1 104 ACETYLATION.  
 DE FT DOMAIN 105 412 HEAD.  
 DE FT DOMAIN 413 >454 ROD.  
 DE FT DOMAIN 105 136 TAIL.  
 DE FT DOMAIN 137 149 COIL 1A.  
 DE FT DOMAIN 150 248 LINKER 1.  
 DE FT DOMAIN 249 265 COIL 1B.  
 DE FT DOMAIN 266 287 LINKER 12.  
 DE FT DOMAIN 288 291 COIL 2A.  
 DE FT DOMAIN 292 411 LINKER 2.  
 DE FT CARBOHYD 47 47 COIL 2B.  
 DE FT CARBOHYD 432 432 GLCNAC (BY SIMILARITY).  
 DE FT NON\_TER 454 454 GLCNAC (BY SIMILARITY).  
 DE SO SEQUENCE 454 AA; 51854 MW; 3DFC58A1 CRC32;  
 DE  
 DE Query Match 9 52.6%; Score 71; DB 1; Length 454;  
 DE Post Local Similarity 50.0%; Pred. No. 2.12e-02;  
 DE Caches 8; Conservative 5; Mismatches 3; Indels 0; Gaps  
 DE Db 232 AFLRSNHEEEVADLLA 247  
 DE QY :!;! : ! : ||| :||:  
 DE OY 5 SFIRTAHDSESEQDLIA 20

| RESULT                                    | 10   | SQ  | SEQUENCE  | 915 AA; 102317 MW; DR95C360 CRC32; |           |
|---|--|---|---|------------------------------------|-----------|
| ID  | NFM_HUMAN  | STANDARD;   | PRT;  | 915 AA.                            |           |
| AC  | P07197;  |   |   |                                    |           |
| DT  | 01-APR-1988 (REL. 08, LAST SEQUENCE UPDATE)  |   |   |                                    |           |
| DT  | 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)  |   |   |                                    |           |
| PRIMATES; CATARRHINI; HOMINIDAE; HOMO; OC | NEUROFILAMENT TRIPLET M PROTEIN (160 KD NEUROFILAMENT PROTEIN) (NF-M).   | DE  | NEFM OR NFM.  | OS                                 |           |
| RN  | HOMO SAPIENS (HUMAN).  | GN  | EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA; |                                    |           |
| RN  | PRIMATES; CATARRHINI; HOMINIDAE; HOMO.   | OC  | RN  |                                    |           |
| [1]                                       | SEQUENCE FROM N.A.   | AC  | PRIMATES; CATARRHINI; HOMINIDAE; HOMO.                        |                                    |           |
| RX  | MEDLINE: 87275853.   | DT  | PRIMATES; CATARRHINI; HOMINIDAE; HOMO.                        |                                    |           |
| RA  | MYERS M. W., LAZZARINI R.A., LEE V.Y., SCHLAEPFER W.W., NELSON D.L.; RL  | DT  | PRIMATES; CATARRHINI; HOMINIDAE; HOMO.                        |                                    |           |
| RR  | "The human mid-size neurofilament subunit: a repeated protein sequence and the relationship of its gene to the intermediate filament gene family";   | DT  | PRIMATES; CATARRHINI; HOMINIDAE; HOMO.                        |                                    |           |
| RL  | EMBO J. 6:1617-1626(1987).   | DE  | PRIMATES; CATARRHINI; HOMINIDAE; HOMO.                        |                                    |           |
| RN  | [2]  | ALPHA-INTERNEVIN (ALPHA-INX).                                       |   |                                    |           |
| RX  | SEQUENCE OF ONE OF THE 13 RESIDUE REPEATS.   | DE  |   |                                    |           |
| RX  | MEDLINE: 88158120.   | ALPHA-INTERNEVIN (ALPHA-INX).                                       |   |                                    |           |
| RA  | LEE V.Y., OTVOS L. JR., CARDEN M.J., HOLLOSI M., DIETZSCHOLD B., RA  | INA   |   |                                    |           |
| RA  | "Identification of the major multiphosphorylation site in mammalian neurofilaments.";  | OS  |   |                                    |           |
| RL  | PROG. NATL. ACAD. SCI. U.S.A. 85:1998-2002(1988).  | MUS   |   |                                    |           |
| CC  | -1-FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M, CC  | MUSCULUS (MOUSE).   |   |                                    |           |
| CC  | AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.   | EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; MAMMALIA; EUTHERIA;       |   |                                    |           |
| CC  | -!- PTM: THERE ARE A NUMBER OF REPEATS OF THE TRIPETIDE K-S-P, NFM IS PROSPHORYLATED ON A NUMBER OF THE SERINES IN THIS MOTIF. IT IS THOUGHT THAT PHOSPHORYLATION OF NFM RESULTS IN THE FORMATION OF INTERFILAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTENANCE OF AXONAL CALIBER.  | OC  |   |                                    |           |
| CC  | -1- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.   | OC  |   |                                    |           |
| CC  | This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement. (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to license@isb-sib.ch). | RA  |   |                                    |           |
| CC  | DR   | DR  |   |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | MGD: MGI:90568; INA.  |   |                                    |           |
| CC  | PIR: A27864; A27864.   | PROSITE; PS00226; IF; 1.  |   |                                    |           |
| CC  | PIR: A30157; A30157.   | PFAM; PF00038; filament; 1.   |   |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | KW  |   |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | INTERMEDIATE FILAMENT; HEPTAD REPEAT PATTERN; COILED COIL; NEURONE. |   |                                    |           |
| CC  | DR   | FT  |   |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DOMAIN  | 1   | 87                                 |           |
| CC  | PIR: A27864; A27864.   | FT  | DOMAIN  | 88                                 | 408       |
| CC  | PIR: A30157; A30157.   | FT  | DOMAIN  | 409                                | 504       |
| CC  | MIM: 152250; MIM: 152250.  | FT  | DOMAIN  | 88                                 | 129       |
| CC  | PROSITE; PS00226; IF; 1.   | FT  | DOMAIN  | 130                                | 142       |
| CC  | DR   | FT  | DOMAIN  | 143                                | 238       |
| CC  | EMBL: Y00677; G35046; -.   | FT  | DOMAIN  | 239                                | 262       |
| CC  | PIR: A27864; A27864.   | FT  | DOMAIN  | 263                                | 408       |
| CC  | PIR: A30157; A30157.   | FT  | DOMAIN  | 449                                | 458       |
| CC  | MIM: 152250; MIM: 152250.  | FT  | DOMAIN  | 504 AA; 55869 MW; 0D788C01 CRC32;  | POLY-GIU. |
| CC  | PROSITE; PS00226; IF; 1.   | SO  | SEQUENCE  |                                    |           |
| CC  | DR   | SO  | SEQUENCE  |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | Query Match   | 51.9%; Score 70; DB 1; Length 504;                            |                                    |           |
| CC  | PIR: A27864; A27864.   | Best Local Similarity   | 50.0%; Pred. No. 3.42e-02;                                    |                                    |           |
| CC  | PIR: A30157; A30157.   | Matches   | 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;            |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | Db  | 225 AFVRQVHDEEVALLA 240                                       |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | Qy  | 5 SFIRTAHDSEVODLIA 20   |                                    |           |
| CC  | DR   | Result  | 12  |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | ID  | A1NK_RAT  |                                    |           |
| CC  | PIR: A27864; A27864.   | AC  | P23565;   |                                    |           |
| CC  | PIR: A30157; A30157.   | DT  | 01-NOV-1991 (REL. 20, CREATED)                                |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DT  | 01-OCT-1995 (REL. 34, LAST SEQUENCE UPDATE)                   |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DT  | 15-JUL-1998 (REL. 36, LAST ANNOTATION UPDATE)                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A27864; A27864.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PIR: A30157; A30157.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | MIM: 152250; MIM: 152250.  | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | PROSITE; PS00226; IF; 1.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | DR   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |
| CC  | EMBL: Y00677; G35046; -.   | DE  | ALPHA-INTERNEVIN (ALPHA-INX).                                 |                                    |           |



CC INTERFILAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTENANCE  
CC OF AXONAL CALIBER.

-!- PTM: PHOSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTIONING

OF THE LARGER NEUROFILAMENT POLYPEPTIDES (NF-M AND NF-H), THE  
LEVELS OF PHOSPHORYLATION BEING ALTERED DEVELOPMENTALLY AND  
COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION.

-!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to license@isb-sib.ch).

DR EMBL: X05640; G297529; -;

DR EMBL: M00481; G200040; -;

DR PIR: S00030; S00030.

DR PROSITE: PS00226; IF; 1.

DR PFAM: PF00038; filament; 1.

KW INTERMEDIATE FILAMENT; HEPTAD REPEAT PATTERN; COILED COIL; NEURONE;

KW PHOSPHORYLATION; GLYCOPROTEIN.

FT INIT\_MET 0 0

FT DOMAIN 1 102 HEAD.

FT DOMAIN 103 410 ROD.

FT DOMAIN 411 848 TAIL.

FT DOMAIN 103 134 COIL 1A.

FT DOMAIN 135 147 LINKER 1.

FT DOMAIN 148 246 COIL 1B.

FT DOMAIN 247 263 LINKER 12.

FT DOMAIN 264 285 COIL 2A.

FT DOMAIN 286 289 LINKER 2.

FT DOMAIN 290 410 COIL 2B.

FT CARBOHYD 430 430 GLCNAC (BY SIMILARITY).

FT CONFLICT 432 539 S -> F (IN REF. 2).

FT CONFLICT 539 540 QA -> RR (IN REF. 2).

SQ 848 AA; 95910 MN; E06A637A CRC32;

Query Match 49.6%; Score 67; DB 1; Length 848;  
Best Local Similarity 50.0%; Pred. No. 1.39e-01;  
Matches 8; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

Db 230 AFLRNNHEEVADLLA 245  
:!: | : || | :| 5 SFIRTAHSEVQDLIA 20

RESULT 15  
ID NFM-CHICK STANDARD; PRT; 857 AA.

ID P16053; STANDARD; PRT; 857 AA.

DT 01-APR-1990 (REL. 1.4, CREATED)

DT 01-APR-1990 (REL. 1.4, LAST SEQUENCE UPDATE)

DT 01-OCT-1996 (REL. 34, LAST ANNOTATION UPDATE)

DE NEUROFILAMENT TRIPLET M PROTEIN (160 KD NEUROFILAMENT PROTEIN) (NF-M).

GN NEFM.

OS GALLUS (CHICKEN).

OC EUKARYOTA; METAZOA; CHORDATA; VERTEBRATA; ARCHOSAURIA; AVES;

OC NEOCORNAE; GALLIFORMES; PHASIANIDAE; PHASIANINAE; GALLUS.

RN [1]

RN SEQUENCE FROM N.A.

RN MEDLINE; 90114973.

RA ZOPF D., DINAYA B., BETZ H., GUNDELINGER E.D.;  
RT "Isolation of the chicken middle-molecular weight neurofilament  
RT (NF-M) gene and characterization of its promoter.";

RL NUCLEIC ACIDS RES. 18:521-529(1990).  
RN [2]

RP SEQUENCE OF 259-857 FROM N.A.

RX MEDLINE; 88112814.

RA ZOPF D., HERMANS-BORGMEYER I., GUNDELINGER E.D., BETZ H.;

RT "Identification of gene products expressed in the developing chick

RT visual system: characterization of a middle-molecular-weight  
RT neurofilament DNA."

RL GENES DEV. 1:699-708(1987).

-!- FUNCTION: NEUROFILAMENTS USUALLY CONTAIN THREE IF PROTEINS: L, M,  
AND H WHICH ARE INVOLVED IN THE MAINTENANCE OF NEURONAL CALIBER.

-!- PTM: THERE ARE A NUMBER OF REPEATS OF THE TRIPETIDE K-S-P, NF-M IS

PHOSPHORYLATED ON A NUMBER OF THE SERINES IN THIS MOTIF. IT IS  
THOUGHT THAT PHOSPHORYLATION OF NFH RESULTS IN THE FORMATION OF  
INTERFILAMENT CROSS BRIDGES THAT ARE IMPORTANT IN THE MAINTENANCE  
OF AXONAL CALIBER.

-!- PTM: PHOSPHORYLATION SEEMS TO PLAY A MAJOR ROLE IN THE FUNCTIONING  
OF THE LARGER NEUROFILAMENT POLYPEPTIDES (NF-M AND NF-H), THE  
LEVELS OF PHOSPHORYLATION BEING ALTERED DEVELOPMENTALLY AND  
COINCIDENT WITH A CHANGE IN THE NEUROFILAMENT FUNCTION.

-!- SIMILARITY: BELONGS TO THE INTERMEDIATE FILAMENT FAMILY.

This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to license@isb-sib.ch).

DR EMBL: X17102; G63689; -;

DR EMBL: X0558; G63686; -;

DR PIR: A21040; A21040.

DR PIR: S08061; S08061.

DR PIR: S15762; S15762.

DR PROSITE: PS00226; IF; 1.

DR PFAM: PF00038; filament; 1.

KW INTERMEDIATE FILAMENT; HEPTAD REPEAT PATTERN; COILED COIL; NEURONE;

KW PHOSPHORYLATION; GLYCOPROTEIN.

FT INIT\_MET 0 0

FT DOMAIN 1 99 HEAD.

FT DOMAIN 99 407 ROD.

FT DOMAIN 131 130 TAIL.

FT DOMAIN 144 143 COIL 1A.

FT DOMAIN 243 242 COIL 1B.

FT DOMAIN 260 259 LINKER 12.

FT DOMAIN 282 281 COIL 2A.

FT DOMAIN 405 405 LINKER 2.

FT CARBOHYD 46 46 GLCNAC (BY SIMILARITY).

FT CARBOHYD 426 426 GLCNAC (BY SIMILARITY).

FT CONFLICT 546 545 G -> R (IN REF. 2).

SQ 857 AA; 95704 MN; 3D05FDD CRC32;

Query Match 49.6%; Score 67; DB 1; Length 857;  
Best Local Similarity 43.8%; Pred. No. 1.39e-01;  
Matches 7; Conservative 5; Mismatches 4; Indels 0; Gaps 0;

Db 226 AFLRNNHEEVADLLA 241  
:!: | : || | :| 5 SFIRTAHSEVQDLIA 20

Search completed: Tue Jan 11 15:42:37 2000  
Job time : 8 secs.







Page  
4

|           |  |
|-----------|--|
| PT        | Classification of lymphoma or lung cancer - on the basis of a point mutation in c-rraf-1 gene  |
| PS        | Example 3; Column 17-20; 26pp; English.  |
| CC        | W1704-W1709 are mutant "versions of the mouse c-rraf-1 protein, derived from four different types of tumour. They were used in a method for classifying a lymphoma or lung cancer, by comparison with the wild-type c-rraf-1 protein sequence derived from different tumours in mice. The method, when used for diagnosing human tumours, involves detecting the presence of a point mutation in a conserved region of the c-rraf-1 gene (codon 533, encoding Ser) derived from lymphoma or lung cancer tissue and classifying the lymphoma or lung cancer as a c-rraf-1 mutation. Associated cancer if one or more point mutations are present. The method is particularly applicable to diagnosis of lung adenocarcinoma.  |
| CC        | Sequence 648 AA;   |
| SQ        | Query Match 41.9%; Score 62; DB 22; Length 648;<br>Best Local Similarity 30.0%; Pred. No. 6.4e+01; Gaps 0;<br>Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;<br>Oy 1 ANDHLSILEAWSNDTPYLH 20  |
| RESULT 9  | ID R22560 standard; Protein; 648 AA.<br>AC R22560;<br>DT 25-AUG-1992 (first entry)<br>DE Mouse mutant c-rraf-1.<br>KW Polymerase chain reaction; tumour; cancer; detection; mutation.<br>OS Mus musculus.<br>FH Key region<br>FT region 62..194 /note= "conserved region (CR) 1"<br>FT region 253..269 /note= "conserved region (CR) 2"<br>FT region 338..627 /note= "conserved region (CR) 3"<br>FT region 513..515 /note= "APE site"<br>FT region 527 /note= "mutation, Phe -> Ser"<br>PR 26-AUG-1988; US-236947.<br>PR 16-SEP-1991; US-759738.<br>PA (USSH ) US DEPT HEALTH & HUMAN.<br>PI Rapp U, Storm S;<br>DR WPI; 92-123681/15.<br>PT Detecting C-RAF-1 genes - by amplifying region of C-RAF-1 gene, analysing prods. and classifying individual having mutation(s) in PT region<br>PS Disclosure; Page 35; 65pp; English.<br>CC The sequence is that of a mutated version of mouse c-rraf-1, the mutation occurs just downstream of the APE site, the mutation is not artificial. The region in which it occurs overlaps an epitope shared by monoclonal antibodies generated against raf (Kolch et al. 1990). This region is a hydrophilic domain, the structure of which is predicted to be altered by this mutation.<br>CC See also R2527 and R22559-R22563.<br>SQ Sequence 648 AA; |
| RESULT 10 | ID R22562 standard; Protein; 648 AA.<br>AC R22562;<br>DT 25-AUG-1992 (first entry)<br>DE Mouse mutant c-rraf-1.<br>KW Polymerase chain reaction; tumour; cancer; detection; mutation.<br>OS Mus musculus.<br>FH Key region<br>FT region 62..194 /note= "conserved region (CR) 1"<br>FT region 253..269 /note= "conserved region (CR) 2"<br>FT region 338..627 /note= "conserved region (CR) 3"<br>FT region 513..515 /note= "APE site"<br>FT region 527 /note= "mutation, Phe -> Ser"<br>PR 16-SEP-1991; 759738.<br>PR 26-AUG-1988; US-236947.<br>PR 16-SEP-1991; US-759738.<br>PA (USSH ) US DEPT HEALTH & HUMAN.<br>PI Rapp U, Storm S;<br>DR WPI; 92-123681/15.   |
| RESULT 11 | ID R22561 standard; Protein; 648 AA.<br>AC R22561;<br>DT 25-AUG-1992 (first entry)<br>DE Mouse mutant c-rraf-1.<br>KW Polymerase chain reaction; tumour; cancer; detection; mutation.<br>OS Mus musculus.<br>FH Key region<br>FT region 62..194 /note= "conserved region (CR) 1"<br>FT region 253..269 /note= "conserved region (CR) 2"<br>FT region 338..627 /note= "conserved region (CR) 3"<br>FT region 513..515 /note= "APE site"<br>FT region 527 /note= "mutation, Asp -> Asn"<br>PR 16-SEP-1991; 759738.<br>PR 26-AUG-1988; US-236947.<br>PR 16-SEP-1991; US-759738.<br>PA (USSH ) US DEPT HEALTH & HUMAN.<br>PI Rapp U, Storm S;<br>DR WPI; 92-123681/15.   |
| RESULT 12 | Query Match 41.9%; Score 62; DB 4; Length 648;<br>Best Local Similarity 30.0%; Pred. No. 6.4e+01; Gaps 0;<br>Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;<br>Oy 1 ANDHLSILEAWSNDTPYLH 20   |

PT Detecting C-Raf-1 genes - by amplifying region of C-Raf-1 gene' in analysing prods. and classifying individual having mutation(s) in region

PT Disclosure: Page 31; 65PP; English.

PS The sequence is that of a mutated version of mouse c-raf-1', the mutation occurs just downstream of the APE site, the mutation is not artificial. The region in which it occurs overlaps an epitope shared by monoclonal antibodies generated against raf (Kolch et al. 1990). This region is a hydrophilic domain, the structure of which is predicted to be altered by this mutation.

CC See also R25277 and R22559-R22563.

SQ Sequence 648 AA;

Query Match 41.9%; Score 62; DB 4; Length 648; Best Local Similarity 30.0%; Pred. No. 6 43e+01; Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Db 413 tkdnlaivtgcggcgslykh 432  
Qy ::|:|:|: | :| :| |

1 ANDHSILEAWSNDNPYLH 20

R T 12 standard; Protein: 648 AA.

ID R25277

AC R25277; (first entry)

PR 16-AUG-1988; US-236947.

DR 16-SEP-1991; US-759138.

PA (USSH ) US DEPT HEALTH & HUMAN.

PI Rapp U, Storm S;

PT DR WPI: 92-133681/15.

N-PSDB: 023459-

PT Detecting C-Raf-1 genes - by amplifying individual having mutation(s) in analysing prods. and classifying

PS Disclosure: Page 46; 65PP; English.

The sequence is that of human c-raf-1, the c-raf-1 gene can be used in methods for detecting and treating a wide range of cancers including lung cancer, T-cell lymphomas, renal cell carcinoma, ovarian carcinoma and mixed parotid gland tumours. The methods consist of amplifying the gene by PCR and then, based on the presence or absence of one or more mutations, individuals at increased risk can be detected and prognosis and treatment determined. See also R22559.

SQ Sequence 648 AA;

Query Match 41.9%; Score 62; DB 4; Length 648; Best Local Similarity 30.0%; Pred. No. 6 43e+01; Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Db 413 tkdnlaivtgcggcgslykh 432  
Qy ::|:|:|: | :| :| |

1 ANDHSILEAWSNDNPYLH 20

RESULT 14 standard; Protein: 648 AA.

ID WI7048

AC WI7048; (first entry)

PR 16-JUN-1997

DE Mutant mouse c-raf 1 protein used in diagnosis of cancer.

KW raf; oncogene; lymphoma; lung cancer; neoplasia; point mutation; conservative region; adenocarcinoma; codon 533; diagnosis; detection.

OS Mus musculus.

PN US511670-A.

PD 08-APR-1997.

PR 26-AUG-1988; 236947.

DR 26-AUG-1988; US-236947.

PS 16-SEP-1991; US-759138.

PR 24-JAN-1994; US-185282.

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PI Rapp U, Storm SM;

DR WPI: 97-225421/20.

PT Classification of lymphoma or lung cancer - on the basis of a point mutation in c-raf-1 gene

PS Example 3: Column 21-24; 25PP; English.

WI7048-WI7049 are mutant versions of the mouse c-raf-1 protein, derived from four different types of tumour. They were used in a method for classifying a lymphoma or lung cancer, by comparison with the wild-type c-raf 1 protein sequence derived from different tumours in mice. The method, when used for diagnosing human tumours, involves detecting the presence of a point mutation in a conserved region of the c-raf-1 gene (codon 533, encoding Ser) derived from lymphoma or lung cancer tissue and classifying the lymphoma or lung cancer as a c-raf-1 mutation. Associated cancer if one or more point mutations are present. The method is particularly applicable to diagnosis of lung adenocarcinoma.

SQ Sequence 648 AA;

Query Match 41.9%; Score 62; DB 22; Length 648; Best Local Similarity 30.0%; Pred. No. 6 43e+01; Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Db 413 tkdnlaivtgcggcgslykh 432  
Qy ::|:|:|: | :| :| |

1 ANDHSILEAWSNDNPYLH 20

**RESULT** 15  
**ID** W17049 standard; Protein; 648 AA.  
**AC** W17049;  
**DT** 16-JUL-1997 (first entry)  
**DE** Mutant mouse c-raf-1 protein used in diagnosis of cancer.  
**KW** raf; oncogene; lymphoma; lung cancer; neoplasia; point mutation; conserved region; adenocarcinoma; codon 533; diagnosis; detection.  
**OS** Mus musculus.  
**PN** US5618670-A.  
**PD** 08-APR-1997.  
**PF** 26-AUG-1988; 236947.  
**PR** 16-SEP-1991; US-759738.  
**PA** (USSH ) US DEPT HEALTH & HUMAN SERVICES.  
**PI** Rapp UR, Storm SM;  
**DR** WPI; 97-225421/20.  
**PT** Classification of lymphoma or lung cancer - on the basis of a point mutation in c-raf-1 gene  
**Example** 3; Column 23-28; 26pp; English.  
**CC** W17046-W17049 are mutant versions of the mouse c-raf-1 protein, derived from four different types of tumour. They were used in a method for classifying a lymphoma or lung cancer, by comparison with the wild-type c-raf-1 protein sequence derived from different tumours in mice. The method, when used for diagnosing human tumours, involves detecting the presence of a point mutation in a conserved region of the c-raf-1 gene (codon 533, encoding Ser - this mutant version has Phe at this position) derived from lymphoma or lung cancer tissue and classifying the lymphoma or lung cancer as a c-raf-1 mutation-associated cancer if one or more point mutations are present. The method is particularly applicable to diagnosis of lung adenocarcinoma.  
**CC** Sequence 648 AA;  
**SQ**

| Query   | Match            | Score | DB         | Length   |
|---------|------------------|-------|------------|----------|
| Best    | Local Similarity | 41.9% | 22         | 648      |
| Matches |                  | 30.0% | Pred. No.  | 6.43e+01 |
|         | 6;               |       | Mismatches | 7;       |
|         | Conservative     |       | Indels     | 0;       |
|         |                  |       | Gaps       | 0;       |

```

Db      413 tkdnlaivtqwgcegsslykh 432
       :|:!:|: | :| |:
Qy      1 ANDHLSILEAWSNDTFLKH 20

```

Search completed: Tue Jan 11 15:36:37 2000  
Job time : 37 secs.





Qy 1 ANDHLSILEAWSDNDTPY 18  
 AC xxxxxx  
 XX DT  
 XX DE Sequence 12, Application US/08909984A  
 XX Patent No. 5747275  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Rubin, Gerry M.  
 CC APPLICANT: Therrien, Marc  
 CC APPLICANT: Chang, Henry C.  
 CC APPLICANT: Wasserman, David A.  
 CC TITLE OF INVENTION: Protein Kinase Required for Ras  
 CC NUMBER OF SEQUENCES: 114  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
 CC STREET: 268 BUSH STREET, SUITE 3200  
 CC CITY: SAN FRANCISCO  
 CC STATE: CALIFORNIA  
 CC COUNTRY: USA  
 CC ZIP: 94104  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: FLOPPY, disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: Patentin Release #1.0, Version #1.30  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/07/857,224B  
 CC FILING DATE: 03/25/92  
 CC CLASSIFICATION: 436  
 CC PRIOR APPLICATION DATA: none  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (International) 41 1 632 2830  
 CC TELEFAX: (International) 41 1 262 2437  
 CC INFORMATION FOR SEQ ID NO: 42:  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 267  
 CC TYPE: amino acid  
 CC TOPOGY: linear  
 CC MOLECULE TYPE:  
 CC DESCRIPTION: protein  
 CC ORGANISM: *Saccharomyces cerevisiae*  
 CC FEATURE: Protein kinase; Table 8 Column 48  
 CC PUBLIC "-//IUB/IUBMB//DTD PROTEIN//EN//EN" 1.0  
 CC AUTHORS: Hanks, S. K.  
 CC AUTORS: Quinn, A. M.  
 CC AUTORS: Hunter, T.  
 CC TITLE: The protein kinase family  
 CC JOURNAL: Science  
 CC VOLUME: 241  
 CC PAGES: 42-52  
 CC DATE: 1988  
 SQ SEQUENCE 267 AA; 30518 MW; 382200 CN;  
 Query Match 41.9%; Score 62; DB 1; Length 315;  
 Best Local Similarity 30.0%; Pred. No. 3.00e+01;  
 Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;  
 Db 68 TKDNLAIVTQCEGSSLYKH 99  
 Qy 1 ANDHLSILEAWSDNDTPYLH 20  
 Db 68 TKDNLAIVTQCEGSSLYKH 99  
 Qy 1 ANDHLSILEAWSDNDTPYLH 20  
 DE Sequence 12, Application US/08909983  
 XX Sequence 12, Application US/08909983  
 CC Patent No. 5747288  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Rubin, Gerry M.  
 CC APPLICANT: Therrien, Marc  
 CC APPLICANT: Chang, Henry C.

CC APPLICANT: Karim, Felix D.  
 CC APPLICANT: Wasserman, David A.  
 CC TITLE OF INVENTION: A No. 574288el Protein Kinase Required for Ras  
 CC NUMBER OF SEQUENCES: 12  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: SCIENCE & TECHNOLOGY LAW GROUP  
 CC STREET: 268 BUSH STREET, SUITE 3200  
 CC CITY: SAN FRANCISCO  
 CC STATE: CALIFORNIA  
 CC COUNTRY: USA  
 CC ZIP: 94104  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/909,983  
 CC FILING DATE: 12-JUN-1997  
 CC CLASSIFICATION: 435  
 CC PRIOR APPLICATION DATA:  
 CC APPLICATION NUMBER: 08/571,758  
 CC FILING DATE:  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: OSMAN, RICHARD A.  
 CC REGISTRATION NUMBER: 36,627  
 CC REFERENCE/DOCKET NUMBER: B96-010  
 CC TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (415) 343-4341  
 CC TELEFAX: (415) 343-4342  
 CC SEQUENCE CHARACTERISTICS:  
 CC LENGTH: 315 amino acids  
 CC TYPE: amino acid  
 CC STRANDEDNESS: not relevant  
 CC MOLECULE TYPE: Peptide  
 CC SEQUENCE 315 AA; 35945 MW; 526392 CN;  
 SQ Query Match 41.9%; Score 62; DB 1; Length 315;  
 Best Local Similarity 30.0%; Pred. No. 3.00e+01;  
 Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;  
 Db 80 TKDNLAIVTQCEGSSLYKH 99  
 Q 1 ANDHLSILEAWSNDNPYLH 20  
 RESULT 7 ID US-08-571-758-12 STANDARD; PRT; 315 AA.  
 DE XX AC XXXXXX  
 XX DT  
 XX Sequence 5, Application US/08276151  
 CC Sequence 5, Application US/08276151  
 CC Patent No. 5557719  
 CC GENERAL INFORMATION:  
 CC APPLICANT: Freed, Ellen  
 CC APPLICANT: Ruggeri, Rosamarie  
 CC TITLE OF INVENTION: Interaction of raf-1 and 14-3-3 Proteins  
 CC NUMBER OF SEQUENCES: 9  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Cooley Godward et al.  
 CC STREET: Five Palo Alto Square  
 CC CITY: Palo Alto  
 CC STATE: CA  
 CC COUNTRY: USA  
 CC ZIP: 94036  
 CC COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.25  
 CC CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/276,151  
 CC FILING DATE: 14-JUL-1994  
 CC CLASSIFICATION: 530  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Torchia, Ph.D., Timothy E.  
 CC REGISTRATION NUMBER: 36,700  
 CC REFERENCE/DOCKET NUMBER: ONYX-005/000US

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 843-5811  
 TELEFAX: (415) 857-0663  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 346 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 HYPOTHETICAL: NO  
 FRAGMENT TYPE: C-terminal  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapiens  
 SQ   SEQUENCE 346 AA; 39386 MW; 635811 CN;

Query Match 41.9%; Score 62; DB 1; Length 346;  
 Best Local Similarity 30.0%; Pred. No. 3 00e+01;  
 Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Db 111 TKDNIAIVTQCEGGSLYKH 130  
 Q :||:||:| :||| :|||  
 1 ANDHLSILEAWSNDNPYLRH 20

RESULT 9  
 ID US-08-185-282-12   STANDARD;   PRT; 648 AA.  
 XX Sequence 12, Application US/08185282  
 XX Sequence 12, Application US/08185282  
 AC Patent No. 5618670  
 XX GENERAL INFORMATION:  
 DT APPLICANT: Rapp, Ulf R.  
 CC APPLICANT: Storm, Stephen M.  
 CC TITLE OF INVENTION: DETECTION METHOD FOR C-RAF-1 GENES  
 CC NUMBER OF SEQUENCES: 12  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: CUSHMAN, DARBY & CUSHMAN  
 CC STREET: 1615 L Street, N.W.  
 CC CITY: Washington  
 CC STATE: D.C.  
 CC COUNTRY: USA  
 ZIP: 20036-5501  
 COMPUTER READABLE FORM:  
 CC COMPUTER TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, version #1.25  
 CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/185,282  
 CC FILING DATE:  
 CC CLASSIFICATION: 424  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Young, Leona G.  
 CC REGISTRATION NUMBER: 37,266  
 CC REFERENCE/DOCKET NUMBER: 25590-0100  
 TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (404) 818-3700  
 CC TELEFAX: (404) 818-3799  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 648 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 AMI-SENSE: NO  
 FRAGMENT TYPE: N-terminal  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapiens  
 SQ   SEQUENCE 648 AA; 73051 MW; 2254837 CN;

Query Match 41.9%; Score 62; DB 2; Length 648;  
 Best Local Similarity 30.0%; Pred. No. 3 00e+01;  
 Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Db 413 TKDNIAIVTQCEGGSLYKH 432  
 QY 1 ANDHLSILEAWSNDNPYLRH 20

RESULT 10  
 ID US-08-886-751A-6   STANDARD;   PRT; 648 AA.  
 XX Sequence 6, Application US/08886751A  
 XX Sequence 6, Application US/08886751A  
 AC Patent No. 5895783  
 XX GENERAL INFORMATION:  
 CC APPLICANT: Yoo, Tai-June  
 CC APPLICANT: Cheng, Kuang-Chuan  
 CC TITLE OF INVENTION: Autoimmune Inner Ear Disease Antigen and Diagnostic Assay  
 CC NUMBER OF SEQUENCES: 6  
 CC CORRESPONDENCE ADDRESS:  
 CC ADDRESSEE: Jones & Askew  
 CC STREET: 191 Peachtree Street, 37th Floor  
 CC CITY: Atlanta  
 CC STATE: Georgia  
 CC COUNTRY: U.S.A.  
 ZIP: 30303  
 COMPUTER READABLE FORM:  
 CC MEDIUM TYPE: Floppy disk  
 CC COMPUTER: IBM PC compatible  
 CC OPERATING SYSTEM: PC-DOS/MS-DOS  
 CC SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 CC APPLICATION NUMBER: US/08/886,751A  
 CC FILING DATE:  
 CC CLASSIFICATION: 424  
 CC ATTORNEY/AGENT INFORMATION:  
 CC NAME: Young, Leona G.  
 CC REGISTRATION NUMBER: 37,266  
 CC REFERENCE/DOCKET NUMBER: 25590-0100  
 TELECOMMUNICATION INFORMATION:  
 CC TELEPHONE: (404) 818-3700  
 CC TELEFAX: (404) 818-3799  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 648 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 HYPOTHETICAL: NO  
 AMI-SENSE: NO  
 FRAGMENT TYPE: N-terminal  
 ORIGINAL SOURCE:  
 ORGANISM: Homo sapiens  
 SQ   SEQUENCE 648 AA; 73051 MW; 2254837 CN;

Query Match 41.9%; Score 62; DB 2; Length 648;  
 Best Local Similarity 30.0%; Pred. No. 3 00e+01;  
 Matches 6; Conservative 7; Mismatches 7; Indels 0; Gaps 0;

Db 413 TKDNIAIVTQCEGGSLYKH 432  
 QY 1 ANDHLSILEAWSNDNPYLRH 20

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (671) 4627 CUSH  
 INFORMATION FOR SEQ ID NO: 12:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 648 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein





Thu Jan 13 09:13:31 2000

US-09-290-049-1.rai

Search completed: Tue Jan 11 15:39:05 2000  
Job time : 8 secs.